

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2001, 10:41:29 ; Search time 1719.24 Seconds  
(without alignments)  
8257.465 Million cell updates/sec

Title: US-09-376-633-1

Perfect score: 1029

Sequence: 1 ATGGATATATGCAATCAACT.....TAACCTTACGAAGAAGTCA 1029

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13168883 seqs, 6898233199 residues

Total number of hits satisfying chosen parameters: 26337766

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents\_NA\_Main:\*

1: /cgn2\_6/ptodata/2/pna/US081\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/pna/US06\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/pna/US07\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/pna/US080\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/pna/US081\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/pna/US082\_COMB.seq.\*  
7: /cgn2\_6/ptodata/2/pna/US083\_COMB.seq.\*  
8: /cgn2\_6/ptodata/2/pna/US084\_COMB.seq.\*  
9: /cgn2\_6/ptodata/2/pna/US085\_COMB.seq.\*  
10: /cgn2\_6/ptodata/2/pna/US086\_COMB.seq.\*  
11: /cgn2\_6/ptodata/2/pna/US087\_COMB.seq.\*  
12: /cgn2\_6/ptodata/2/pna/US088\_COMB.seq.\*  
13: /cgn2\_6/ptodata/2/pna/US089\_COMB.seq.\*  
14: /cgn2\_6/ptodata/2/pna/US090\_COMB.seq.\*  
15: /cgn2\_6/ptodata/2/pna/US091\_COMB.seq.\*  
16: /cgn2\_6/ptodata/2/pna/US092\_COMB.seq.\*  
17: /cgn2\_6/ptodata/2/pna/US093\_COMB.seq.\*  
18: /cgn2\_6/ptodata/2/pna/US094\_COMB.seq.\*  
19: /cgn2\_6/ptodata/2/pna/US095A\_COMB.seq.\*  
20: /cgn2\_6/ptodata/2/pna/US095B\_COMB.seq.\*  
21: /cgn2\_6/ptodata/2/pna/US095C\_COMB.seq.\*  
22: /cgn2\_6/ptodata/2/pna/US095D\_COMB.seq.\*  
23: /cgn2\_6/ptodata/2/pna/US096A\_COMB.seq.\*  
24: /cgn2\_6/ptodata/2/pna/US096B\_COMB.seq.\*  
25: /cgn2\_6/ptodata/2/pna/US096C\_COMB.seq.\*  
26: /cgn2\_6/ptodata/2/pna/US096D\_COMB.seq.\*  
27: /cgn2\_6/ptodata/2/pna/US096E\_COMB.seq.\*  
28: /cgn2\_6/ptodata/2/pna/US097A\_COMB.seq.\*  
29: /cgn2\_6/ptodata/2/pna/US097B\_COMB.seq.\*  
30: /cgn2\_6/ptodata/2/pna/US097C\_COMB.seq.\*  
31: /cgn2\_6/ptodata/2/pna/US098\_COMB.seq.\*  
32: /cgn2\_6/ptodata/2/pna/US098A\_COMB.seq.\*  
33: /cgn2\_6/ptodata/2/pna/US098B\_COMB.seq.\*  
34: /cgn2\_6/ptodata/2/pna/US098C\_COMB.seq.\*  
35: /cgn2\_6/ptodata/2/pna/US098D\_COMB.seq.\*  
36: /cgn2\_6/ptodata/2/pna/US098E\_COMB.seq.\*  
37: /cgn2\_6/ptodata/2/pna/US098F\_COMB.seq.\*  
38: /cgn2\_6/ptodata/2/pna/US098G\_COMB.seq.\*  
39: /cgn2\_6/ptodata/2/pna/US098H\_COMB.seq.\*  
40: /cgn2\_6/ptodata/2/pna/US098I\_COMB.seq.\*  
41: /cgn2\_6/ptodata/2/pna/US098J\_COMB.seq.\*  
42: /cgn2\_6/ptodata/2/pna/US098K\_COMB.seq.\*  
43: /cgn2\_6/ptodata/2/pna/US098L\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1029	100.0	1029	1 PCT-US98-24852-1	Sequence 1, Appl1
2	1029	100.0	1029	17 US-09-369-760-1	Sequence 1, Appl1
3	1029	100.0	1029	17 US-09-376-633-1	Sequence 1, Appl1
4	1029	100.0	3336	1 PCT-US98-24852-2	Sequence 7, Appl1
5	1029	100.0	3336	1 PCT-US98-24852-3	Sequence 7, Appl1
6	1029	100.0	3336	1 PCT-US98-24852-4	Sequence 7, Appl1
7	1029	100.0	3336	1 PCT-US98-24852-5	Sequence 7, Appl1
8	1029	100.0	3336	1 PCT-US98-24852-6	Sequence 7, Appl1
9	1029	100.0	3336	13 US-08-977-553-5	Sequence 7, Appl1
10	1029	100.0	3336	13 US-08-977-866-7	Sequence 7, Appl1
11	1029	100.0	3336	16 US-09-227-806-7	Sequence 7, Appl1
12	1029	100.0	3336	17 US-09-369-700-7	Sequence 7, Appl1
13	1029	100.0	3336	17 US-09-376-633-7	Sequence 7, Appl1
14	1029	100.0	3336	18 US-09-492-479-7	Sequence 7, Appl1
15	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
16	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
17	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
18	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
19	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
20	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
21	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
22	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
23	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
24	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
25	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
26	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
27	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
28	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
29	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
30	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
31	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
32	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
33	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
34	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
35	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
36	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
37	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
38	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
39	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
40	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
41	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
42	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
43	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
44	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap
45	1029	100.0	3336	13 US-08-827-356-1376	Sequence 1376, Ap

## ALIGNMENTS

RESULT 1  
PCT-US98-24857-1  
Sequence 1, Application PC/TUS9824857A  
GENERAL INFORMATION:  
APPLICANT: SmithKline Beecham Corporation  
TITLE OF INVENTION: Novel ribg  
FILE REFERENCE: P50444-9  
CURRENT APPLICATION NUMBER: PCT/US98/24857A  
CURRENT FILING DATE: 1998-11-23  
EARLIER APPLICATION NUMBER: 08/978,456  
EARLIER FILING DATE: 1997-11-25  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 1029  
TYPE: DNA  
ORGANISM: Staphylococcus aureus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(1026)  
PCT-US98-24857-1

Query Match 100.0%; Score 1029; DB 1; Length 1029;  
Best Local Similarity 100.0%; Pred. No. 6.e-224;  
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGATTTATGCGATTCACACTTGCATAATATGTACAAAGTCACAAAGCTTTATTCACCCC 60
DB 1 atgattatgcatcacttgcataatattgtacaaagtcacaaagctttatccccc 60
QY 61 GTTGGCGCTGTGTGATTAAATGAAGTAGAGATTGTTGATTGTCACACTTGAAGAAA 120
DB 61 gttggcgctgtgtgattaaatgaagtagagattgttgattgtgacacttggaaaaa 120
QY 121 GGTGCAAGCATCGCAGAGCTTACGACTTGAATGCGCAACAACAANAATGCTGAAGTGGC 180
DB 121 ggtgcaagcatcgcagagcttacgacttgaatgcgcaacaacaanaatgctgaagtggc 180
QY 181 ACGATTATATAGCGTTAGAGCCATGATGTCATTTGTTCAACCCACCTGTTTAC 240
DB 181 acgattatattagcgttagagccatgatgtcatTTTGTTCACCCACCTGTTTAC 240
QY 241 AAAATTATTTGATTGTAAGATACCAAAAGTAGTATACGCAACAACAAGACATTCGTTAGC 300
DB 241 aaaattatttattgtaagataccaaaagtagtattacgcaacaacaagacattcgttagac 300
QY 301 ACACATGCTGATGACAGCTTACGGGCTCAGCGTATTGAGCTTGAATGCGTTGATGATGAA 360
DB 301 acacatgctgatgacagcttacgggctcagcgattgagcttgaatgcgttgatgataa 360
QY 361 CGGCGATCACAATTATACCAAGCTTTTAAAGCAAAAGCAACCAATGTCACAAAT 420
DB 361 cggcgatcacaaattataccagctttttaaagcaaaagcaaaccaatgaccacaatt 420
QY 421 ACAGTGAAGATATCTGCAAGTTTAGATGCTAAACAAGCAGCAATGATGAGCAAAAGTCAA 480
DB 421 acagtgaagatattctgcaagtttagatgcttaacaagcagcaatgatgagcaaaaagtcaa 480
QY 481 TGGATTACTACCAAGAGGTTTACACAGATGCTATTAAGTTAAGACACAGCAGCGCA 540
DB 481 tggattactaccaagaggtttaacaagatgctataagtttaagacatgcagcagcga 540
QY 541 GTGTTAAGTGAAGAGAGTACAGTGAATAGATGATCCACATATATACAGTATTCGA 600
DB 541 gtgttaagtgaagagagtacagtgtaattagatgacacatactacacagatltcaa 600
QY 601 GATGCAAAAAACCCCTATATAAAGTAATATATGTTAAGTCTGGGAATATTCATTTATCAG 660
DB 601 gatgcaaaaaacccctataaagaatattgtctaagctcgggaatatattcaattatcag 660

```

```

QY 661 CAAATTTATCAAGATGATCAACACCAATTGGATATATCTGAAAAATCCAAATTTACA 720
DB 661 caaatttatcaagatgatacaacacaaattggatatactgaaataccaaattaca 720
QY 721 AGCATTCAAACACATATGAAATTTATTTACTGAGAGCTCTGTGATTTAAACAACTTCT 780
DB 721 agcattcAAACACATATGAAATTTATTTACTGAGAGCTCTGTGATTTAAACAACTTCT 780
QY 781 CACAATTTATATAAAGAGAGAGTGAACCTTGTAGTCAGAGCGAGTCCAACTACT 840
DB 781 cacaatttatataaagagagagtggaacttgcagtcagcagcagcagcagcact 840
QY 841 TCAGAAATTCATCTATTAATATAGATGATTAATTCCTATTAATCCCGAAATTAAT 900
DB 841 tcagaattcatctattatataagatgaatttctctatattgccccgaattaat 900
QY 901 GCGGATCTGGAATTTATCAATTTATCAAAACAAATGATGATGATGATGATGATGATG 960
DB 901 gcggatctggaatTTATCAATTTATCAAAACAAATGATGATGATGATGATGATGATG 960
QY 961 AACCAATTTGAATTTGATTCAGTTCGAGTTATTAATCAAAATGTTAAATTAATTAAT 1020
DB 961 aaccatttgaattgttcatcgcagttatataatcaaaaatgttaattaaatttaca 1020
QY 1021 AAGAGTGA 1029
DB 1021 aagaagtga 1029

```

RESULT 2  
US-09-369-700-1  
Sequence 1, Application US/09369700  
GENERAL INFORMATION:  
APPLICANT: Palmer, Leslie M.  
TITLE OF INVENTION: Novel ribg  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESS: Dechert Price & Rhoads  
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
CITY: Philadelphia  
STATE: PA  
COUNTRY: US  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/369,700  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/978,456  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Dickinson, Todd Q  
REGISTRATION NUMBER: 28,354  
REFERENCE/DOCKET NUMBER: P50444-9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-994-2252  
TELEFAX: 215-994-2222  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1029 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-09-369-700-1

Query Match	100.0%;	Score 1029;	DB 17;	Length 1029;
Best Local Similarity	100.0%;	Pred. No. 6.6e-224;		
Matches 1029;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	ATGATTTATGCGATTCAACTTGGCAATATPVGSTGTCAGAGGTCGCAACGAGGTTTAATCCACC	60
Dp	1	ATGATTTATGCGATTCAACTTGGCAATATPVGSTGTCAGAGGTCGCAACGAGGTTTAATCCACC	60
QY	61	GTTCGGCCTGTGTGATTTAATGAAGGTAGGATTTGGTGTATTTGGTGCACACTTGTGAGAAA	120
Dp	61	GTTCGGCCTGTGTGATTTAATGAAGGTAGGATTTGGTGTATTTGGTGCACACTTGTGAGAAA	120
QY	121	GGTGACACGATCGCGGAGGTTTCAAGCACTTGATATGGCACAACAAAATGCTGAAGTGG	180
Dp	121	GGTGACACGATCGCGGAGGTTTCAAGCACTTGATATGGCACAACAAAATGCTGAAGTGG	180
QY	181	ACGATTTATTTATGTTAAGTGAAGCCATGATGATCTTTGGTTCACACACACCCGTGTTCAC	240
Dp	181	ACGATTTATTTATGTTAAGTGAAGCCATGATGATCTTTGGTTCACACACACCCGTGTTCAC	240
QY	241	AAATTTATGTTGTGAATAGCAAAAAGTAGTATCGCAACAAAAGACAAATTCGTTAAC	300
Dp	241	AAATTTATGTTGTGAATAGCAAAAAGTAGTATCGCAACAAAAGACAAATTCGTTAAC	300
QY	301	ACACATGCTGATGAGACGTTACGGGCTCACGGTATTGAGGTTGAATGCGTTGATGATCAA	360
Dp	301	ACACATGCTGATGAGACGTTACGGGCTCACGGTATTGAGGTTGAATGCGTTGATGATCAA	360
QY	361	CGGGCATCACAAATTAATCCAAAGACTTTTTTAAAGCAAAAGCAACACTGCGACAAATTT	420
Dp	361	CGGGCATCACAAATTAATCCAAAGACTTTTTTAAAGCAAAAGCAACACTGCGACAAATTT	420
QY	421	ACAGTGAAGTAGTCTGCAAGTTAGTGGTAAACAAGCAATGATTAATGGACAAAGTCAA	480
Dp	421	ACAGTGAAGTAGTCTGCAAGTTAGTGGTAAACAAGCAATGATTAATGGACAAAGTCAA	480
QY	481	TGGATTTCTAACAAGAGGTTAAACAAGATGCTATAATTTAAGCATTCGACACAGCCA	540
Dp	481	TGGATTTCTAACAAGAGGTTAAACAAGATGCTATAATTTAAGCATTCGACACAGCCA	540
QY	541	GTTGTTAACTGGAGACGCTACAGTTGGAATTAGATGATCCAAATTAATCAACAGTATTCGA	600
Dp	541	GTTGTTAACTGGAGACGCTACAGTTGGAATTAGATGATCCAAATTAATCAACAGTATTCGA	600
QY	601	GATGGAAGAAAAACCTATATAAGTAATATGTCTAAGCTGGGAAATATTCATTTAATCAG	660
Dp	601	GATGGAAGAAAAACCTATATAAGTAATATGTCTAAGCTGGGAAATATTCATTTAATCAG	660
QY	661	CAATTTATCAAGTGAATCAACCAATTTGGATTAATACGAAATCCAAATTTAACA	720
Dp	661	CAATTTATCAAGTGAATCAACCAATTTGGATTAATACGAAATCCAAATTTAACA	720
QY	721	AGCAATCAAAACATATTTAAATTTACTTGAAGTCTGTGATTTAAACAATTCCTT	780
Dp	721	AGCAATCAAAACATATTTAAATTTACTTGAAGTCTGTGATTTAAACAATTCCTT	780
QY	781	CACATTTATTAATAAGAGAGATTGGAACCTTGTAGTGAGGCGAGTCCAACTACT	840
Dp	781	CACATTTATTAATAAGAGAGATTGGAACCTTGTAGTGAGGCGAGTCCAACTACT	840
QY	841	TCAAAATTCCTCATCTTTATATAGATGAATTTATTTCTATATTAATGCCCGAAATTAAT	900
Dp	841	TCAAAATTCCTCATCTTTATATAGATGAATTTATTTCTATATTAATGCCCGAAATTAAT	900
QY	901	GGCGGATCTGGAATTTATCAATTTTATCAACAATGATGATGATGAGATACAGATCG	960
Dp	901	GGCGGATCTGGAATTTATCAATTTTATCAACAATGATGATGATGAGATACAGATCG	960
QY	961	AACCAATTTGAAATTTGTTCAATTCGAGATTATAACCAAAAGTTAAATTAACCTTTACGA	1020
Dp	961	AACCAATTTGAAATTTGTTCAATTCGAGATTATAACCAAAAGTTAAATTAACCTTTACGA	1020
QY	1021	AAGAAGTGA	1029

Db 1021 AAGAGCTGA 1029

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	100.0%;	Score 1029;	DB 17;	Length 1029;
Best Local Similarity	100.0%;	Pred. No. 6.6e-224;		
Matches 1029; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

Oy	1	ATGAGTTATGCGATTCACACTGCAAAATATGGTACAAAGTCAAAACAGGTTAAATCCACC	60
Db	1	ATGGAATTATGGGATTTCAACACTTGCAAAATATGGTACAAAGTCAAAACAGGTTAAATCCACC	60
Oy	61	GTGGCGCTGTTTGTAAGTAAATGAAGGTAGCAATTTGGTATGTCACACTTGTAGAAAA	120
Db	61	GTGGCGCTGTTTGTAAGTAAATGAAGGTAGCAATTTGGTATGTCACACTTGTAGAAAA	120
Oy	121	GGTGCAAGCATGCGGAGAGTTCAAGCACACTTCAATATGCGACAAACAAATGCTGAAGTGGCG	180
Db	121	GGTGCAAGCATGCGGAGAGTTCAAGCACACTTCAATATGCGACAAACAAATGCTGAAGTGGCG	180
Oy	181	ACGATTATATTACGTTAGAGACCACTGATGCAATTTTGGTCAACACCACTCGTGTAAAC	240
Db	181	ACGATTATATTACGTTAGAGACCACTGATGCAATTTTGGTCAACACCACTCGTGTAAAC	240
Oy	241	AAAATTTATGATTGTAAGATATGCAAAAAGTACTATACGACACAAAGAACAATTCGTTAGAC	300
Db	241	AAAATTTATGATTGTAAGATATGCAAAAAGTACTATATACGACACAAAGAACAATTCGTTAGAC	300
Oy	301	ACACATGGTGTATGAGACGTTACGGGGCTACCGGTATTTGAGGTGTAATCCGTTAGATGAAC	360
Db	301	ACACATGGTGTATGAGACGTTACGGGGCTACCGGTATTTGAGGTGTAATCCGTTAGATGAAC	360

Db 301 ACACATGATGAGACGTTACGGCTCACGGTATTGAGGTGATGCGTTGATGATGA 360  
Qy 361 CGGGCATCACAATTATVACCAAGACTTTTAAAGCAAAAGCAAGCACTGCCCAAAAT 420  
Db 361 CGGGCATCACAATTATVACCAAGACTTTTAAAGCAAAAGCAAGCACTGCCCAAAAT 420  
Qy 421 ACAGTGAAGATATCTGCAAGTTAGATGTTAAACGAAGATATATGACAAAGTCAA 480  
Db 421 ACAGTGAAGATATCTGCAAGTTAGATGTTAAACGAAGATATATGACAAAGTCAA 480  
Qy 481 TGGATTACTAACAAAGAGTTAAACAGATGCTATAGTTAAAGATGACGACGCGCA 540  
Db 481 TGGATTACTAACAAAGAGTTAAACAGATGCTATAGTTAAAGATGACGACGCGCA 540  
Qy 541 GTGTTAACTGGAAGACGTACGTGATTAATAGATGATCCAAATATCTACAGCTATTCAA 600  
Db 541 GTGTTAACTGGAAGACGTACGTGATTAATAGATGATCCAAATATCTACAGCTATTCAA 600  
Qy 601 GATGGAAGAAACCCCTATTAAGTATATTTGCTAGTGGGAAATATTTAATTCAG 660  
Db 601 GATGGAAGAAACCCCTATTAAGTATATTTGCTAGTGGGAAATATTTAATTCAG 660  
Qy 661 CAAATTTATCAAGATGATCAACACCAATTTGGATATATCTGAAATCCAAATTTAACA 720  
Db 661 CAAATTTATCAAGATGATCAACACCAATTTGGATATATCTGAAATCCAAATTTAACA 720  
Qy 721 ACCAATCAACAACATATTTGAATTTATTTACTGAACTCTTGATTTAAACAACAAATCTT 780  
Db 721 ACCAATCAACAACATATTTGAATTTATTTACTGAACTCTTGATTTAAACAACAAATCTT 780  
Qy 781 CCAATTTATTAAGAGAGAGTGTGAACTTGTGAGAGGAGGAGTCCAACTACT 840  
Db 781 CCAATTTATTAAGAGAGAGTGTGAACTTGTGAGAGGAGGAGTCCAACTACT 840  
Qy 841 TCAGATTTCCATCTATATATATAGATGATTTATCTGATTTATCCCGCAATTTAAT 900  
Db 841 TCAGATTTCCATCTATATATATAGATGATTTATCTGATTTATCCCGCAATTTAAT 900  
Qy 901 GCGGATCTGGAATATATCAATTTATCAACAAGATGATGATGATCCAGATGCG 960  
Db 901 GCGGATCTGGAATATATCAATTTATCAACAAGATGATGATGATGATCCAGATGCG 960  
Qy 961 AACCAATTTGAATGTTGATCCGAGTATTAATCAAAATGTTAAATTAATTTACGA 1020  
Db 961 AACCAATTTGAATGTTGATCCGAGTATTAATCAAAATGTTAAATTAATTTACGA 1020  
Qy 1021 AAGAGTGA 1029  
Db 1021 AAGAGTGA 1029

RESULT 4  
PCT-US98-24857-7  
: Sequence 7, Application PC/TUS9824857A  
: GENERAL INFORMATION:  
: APPLICANT: SmithKline Beecham Corporation  
: TITLE OF INVENTION: Novel r1b6  
: FILE REFERENCE: P5044-9  
: CURRENT APPLICATION NUMBER: PCT/US98/24857A  
: CURRENT FILING DATE: 1998-11-23  
: EARLIER APPLICATION NUMBER: 08/978,456  
: EARLIER FILING DATE: 1997-11-25  
: NUMBER OF SEQ ID NOS: 7  
: SOFTWARE: FastSeq for Windows version 3.0  
: SEQ ID NO 7  
: LENGTH: 3336  
: TYPE: DNA  
: ORGANISM: Staphylococcus aureus  
PCT-US98-24857-7

Query Match 100.0%; Score 1029; DB 1; Length 3336;  
Best Local Similarity 100.0%; Pred. No. 9.2e-224;

Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGATTTATGCGATTTCACCTGCAATATGATGACAAAGTCAACAGGTGTTATCCACCC 60  
Db 1 atgattatgcatcttaacttgcgaataatgatacaaggtcaaacagggttataatccacc 60  
Qy 61 GTTGGCGTGTGTTAGTATATGAAAGTAGATTTGTTGTTATTTGTCACACTTGAGAAAA 120  
Db 61 gttggcggtgtgtttagtattatgaaagtagattttgtttgtttgttcacacttgagaaaa 120  
Qy 121 GGTGCAACATGCGGAGGTTCACGACCTGATATGCGCAACCAAAATCTGAAGTGGC 180  
Db 121 ggtgcaacatgcgaggttcacgacctgatatgcgcaaccaaactgaagtgag 180  
Qy 181 ACAGTTTATATACGTTAACGACATGATCTATTTGTTTAAACACCACTGTGTTAAC 240  
Db 181 acagtttatatacgtttaacgacatgattctttttaaagcaaacaccctgtgttaac 240  
Qy 241 AAAATTTATGATTTGATAGTACCAAAAAGTATATACGCAAAAAGCAATTCGTTAGC 300  
Db 241 aaatattatgtatgataagtagcaaaagtagatacgcaaaaagcaattcgttagc 300  
Qy 301 ACACATGATGATGAGACGTTACGGGCTCACGGTATTTGAGTTGAATGCTGTGATGTA 360  
Db 301 acacatgatgatgagacgttacgggctcacggatgttgaatgctgtgatgtga 360  
Qy 361 CGGGCATCACAATTATVACCAAGACTTTTAAAGCAAAAGCAAGCACTGCCCAAAAT 420  
Db 361 cgggcatacacaattatvaccagactttttaaagcaaaaagcaagcaactgcccaaat 420  
Qy 421 ACAGTGAAGATATCTGCAAGTTAGATGTTAAACAAGCAAGATATATGACAAAGTCAA 480  
Db 421 acagtgaagatattctgcaagtttagatggtlaaacaagcgaaatgataatgaaagttcaa 480  
Qy 481 TGGATTACTAACAAAGAGTGTAAACAAGATGCTATATTTAAAGATGACGACGCGCA 540  
Db 481 tggattactaacaaagagtttaacaagatgcttatattttaaagatgacgacgcgca 540  
Qy 541 GTGTTAACTGGAAGACGTACGTGATTAATAGATGATCCAAATATCTACAGTATTCAA 600  
Db 541 gttgtaactggaagacgtacgtgattagatgattcccaatatatactacagattcaa 600  
Qy 601 GATGGAAGAAACCCCTATTAAGTATATTTGCTAGTGGGAAATATTTAATTCAG 660  
Db 601 gatggaagaaacccctattaaagtattttgctagtgggaaatatttttaattcag 660  
Qy 661 CAAATTTATCAAGATGATCAACACCAATTTGGATATATCTGAAATCCAAATTTAACA 720  
Db 661 caaatttatcaagatgatcaacaccaatttggatattatctgaaatccaaatttaca 720  
Qy 721 ACCAATCAACAACATATTTGAATTTATTTACTGAACTCTTGATTTAAACAACAAATCTT 780  
Db 721 accaatcaacaacatatttgaatttattttactgaaactcttgattttaacaacaatctt 780  
Qy 781 CCAATTTATTAAGAGAGAGTGTGAACTTGTGATGAGGAGGAGTCCAACTACT 840  
Db 781 ccaatttatataaagagagtgtaacttgcatagtgagagaggtcccaactact 840  
Qy 841 TCAGATTTCCATCTATATATATAGATGATTTATCTGATTTATCCCGCAATTTAAT 900  
Db 841 tcagatttccatctatattatattagatgatttattctgatttatacccgcaatttaat 900  
Qy 901 GCGGATCTGGAATATATCAATTTATCAACAAGATGATGATGATCCAGATGCG 960  
Db 901 gcggatctggaatattatcaatttatcaacaagatgatgatgattccagatgag 960  
Qy 961 AACCAATTTGAATGTTGATCCGAGTATTAATCAAAATGTTAAATTAATTTACGA 1020  
Db 961 aaccaatttgaatgttgcagattatcaaaatgtaaaatgtaaaatgtaaaatgtaaa 1020  
Qy 1021 AAGAGTGA 1029  
Db 1021 aagaagtga 1029



QY 361 CGGGCATCAAAATTATACCAAGACTTTTAAAGCAAAAGCAAGACTGCCAATAAT 420  
|||||  
Db 361 cgggcatcaaaattataccagactttttaaagcaaaagcaagcaagcaccacaatt 420  
QY 421 ACAGTGAAGATATCTCAAGATTAGTGTAAACAGCGAATGATATGCAAAATCA 480  
|||||  
Db 421 acagtgaagatctcgaagtttagtggtaaacgaagcgaatgataatgacaaatc 480  
QY 481 TGGATTACTAACCAAGAGTTAAACAAGATGTTATAGTTAAGACATGCACGCG 540  
|||||  
Db 481 tggattactaacaaagaggtlaaacagaatgctataaagttaagacatcgaccgagc 540  
QY 541 GGTGTACTGGAAGAGCTCAGTGAATTAAGATGATCCAAATATCTCAGTATTC 600  
|||||  
Db 541 ggtgtactggaagagctcagtggaattagatgacccaataatactacagttac 600  
QY 601 GATGAAAAAACCCCTTAAAGTAATTTGCTAACTGCGAATTTCTATTTAAAT 660  
|||||  
Db 601 gatgaaaaaaccccttaaaagtaatttgcctaaagtcagggaataatcatttaac 660  
QY 661 CAAATTTATCAAGATGATCAACACCAATTTGGATATATCTGAAATCCAAAT 720  
|||||  
Db 661 caaatttatcaagatgaatcaacaccaatttgatataatactgaaatccaaat 720  
QY 721 AGCAATCAACACATTTGAATTTATTTACTGAGTCTTGATTTACCAATTC 780  
|||||  
Db 721 agcaatcaaacacatctgaatattactcgaagtcctggaatttaacaaatctt 780  
QY 781 CACAAATTTATTAAGAGAGAGTTGGAAGTTCCTAGTGGAGCGACGCTCCAA 840  
|||||  
Db 781 cacaattatataaagagagtggaacttctagtcgagcgagtcgaacactact 840  
QY 841 TGAGATTCTCCATCTATATATAGTGAATTTATCTATTTATGCCGGAATTA 900  
|||||  
Db 841 tgaattctccatctattatataagaatgaaatctatctatcccgaaatlaatt 900  
QY 901 GCGGATCTGGAATATATCAATTTTCAACAAATGATGTGATGATACAGATG 960  
|||||  
Db 901 gcggatctggaatataatcttcaacaaatgataatgataatcgaatcg 960  
QY 961 AACCAATTTGAATTTGTTGATTCGAGTATTAATCAAAATGTTAAATTA 1020  
|||||  
Db 961 aaccaattggaattgctatccgagtgatlaataacaaatgtaataacttaaga 1020  
QY 1021 AAGAGTGA 1029  
|||||  
Db 1021 aagaagtga 1029

RESULT 7  
PCT-US98-25096-5  
; Sequence 5, Application PC/TUS9825096  
; GENERAL INFORMATION:  
; APPLICANT: SmithKline Beecham Corporation  
; TITLE OF INVENTION: ribh  
; FILE REFERENCE: GM10123  
; CURRENT APPLICATION NUMBER: PCT/US98/25096  
; CURRENT FILING DATE: 1998-11-24  
; EARLIER APPLICATION NUMBER: 08/977,553  
; EARLIER FILING DATE: 1997-11-25  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 5  
; LENGTH: 3336  
; TYPE: DNA  
; ORGANISM: Staphylococcus aureus  
PCT-US98-25096-5

Query Match 100.0%; Score 1029; DB 1; Length 3336;  
Best Local Similarity 100.0%; Pred. No. 9,26-224;  
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGATTTATGCCATTCACCTTGAATATGTCAGAGTCAACAGAGTGTATATCCACC 60  
|||||  
Db 1 atgatttatgccatttcaccttgaaatattgctacaggtcaaacaggtgtatccacc 60  
QY 61 GTTGGCGCTGTTGTTAGTTATGAGGTAGGATTTGTTGTTGTTGTTGTTGTTG 120  
|||||  
Db 61 gttggcgctgtttagttatgaggttaggattgttggattgttggattgttggatt 120  
QY 121 GTGACAAAGCATGCGAGGTTCAAGCATTGATATGCGCAACAAATGCTGAAG 180  
|||||  
Db 121 gtgacaaagcatgcgaggttcaagcattgatattgcaacaataatgctgaagtt 180  
QY 181 ACATTTATATATACGTTATAGCATGATTTTGGTTCAACACCACTGTTATAC 240  
|||||  
Db 181 acatttatatatcgttatagcatgatTTTTGGTTCAACACCACTGTTATAC 240  
QY 241 AAAATTTGATTTGTAAGTATAGCAAAAGATATACGCAAAAGCAAAATGTTAG 300  
|||||  
Db 241 aaaatttatgatttgaagtataagtaaaagtataacgcaaaaagcaaatcgttag 300  
QY 301 ACACATGTTGATGAGACGTTACGGGCTCAGGATTTGAGTTGAATGCTGATG 360  
|||||  
Db 301 acacatgttgatgagacgtttacgggctcaggatttgaggttgaaatgctgatg 360  
QY 361 CGGGCATCAAAATTATACCAAGACTTTTAAAGCAAAAGCAAGACTGCCAATA 420  
|||||  
Db 361 cgggcatcaaaattataccagactttttaaagcaaaagcaagcaagcaccacaatt 420  
QY 421 ACAGTGAAGATATCTCAAGATTAGTGTAAACAGCGAATGATATGCAAAAGT 480  
|||||  
Db 421 acagtgaagatctcgaagtttagtggtaaacgaagcgaatgataatgacaaatc 480  
QY 481 TGGATTACTAACCAAGAGTTAAACAAGATGTTATAGTTAAGACATGCACGCG 540  
|||||  
Db 481 tggattactaacaaagaggtlaaacagaatgctataaagttaagacatcgacga 540  
QY 541 GGTGTACTGGAAGAGCTCAGTGAATTAAGATGATCCAAATTTACTGAGTAT 600  
|||||  
Db 541 ggtgtactggaagagctcagtggaattagatgacccaatTTACTGAGTAT 600  
QY 601 GATGAAAAAACCCCTTAAAGTAATTTGCTAACTGCGAATTTCTATTTAAAT 660  
|||||  
Db 601 gatgaaaaaaccccttaaaagtaatttgcctaaagtcagggaataatcatttaac 660  
QY 661 GATGAAAAAACCCCTTAAAGTAATTTGCTAACTGCGAATTTCTATTTAAAT 720  
|||||  
Db 661 gatgaaaaaaccccttaaaagtaatttgcctaaagtcagggaataatcatttaac 720  
QY 721 AGCAATCAACACATTTGAATTTATTTACTGAGTCTTGATTTACCAATTC 780  
|||||  
Db 721 agcaatcaaacacatctgaatattactcgaagtcctggaatttaacaaatctt 780  
QY 781 CACAAATTTATTAAGAGAGAGTTGGAAGTTCCTAGTGGAGCGACGCTCCAA 840  
|||||  
Db 781 cacaattatataaagagagtggaacttctagtcgagcgagtcgaacactact 840  
QY 841 TGAGATTCTCCATCTATATATAGTGAATTTATCTATTTATGCCGGAATTA 900  
|||||  
Db 841 tgaattctccatctattatataagaatgaaatctatctatcccgaaatlaatt 900  
QY 901 GCGGATCTGGAATATATCAATTTTCAACAAATGATGTGATGATACAGATG 960  
|||||  
Db 901 gcggatctggaatataatcttcaacaaatgataatgataatcgaatcg 960  
QY 961 AACCAATTTGAATTTGTTGATTCGAGTATTAATCAAAATGTTAAATTA 1020  
|||||  
Db 961 aaccaattggaattgctatccgagtgatlaataacaaatgtaataacttaaga 1020  
QY 1021 AAGAGTGA 1029  
|||||  
Db 1021 aagaagtga 1029

```
RESULT 8
US-08-977-553-5
: Sequence 5, Application US/08977553
: GENERAL INFORMATION:
: APPLICANT: Palmer, Leslie M.
: TITLE OF INVENTION: Novel ribb
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dechart Price & Rhoads
: STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
: City: Philadelphia
: STATE: PA
: COUNTRY: US
: ZIP: 19103
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/977,553
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Dickinson, Todd O
: REGISTRATION NUMBER: 28,354
: REFERENCE/DOCKET NUMBER: GM10123
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 215-994-2252
: TELEFAX: 215-994-2222
: TELEX:
: INFORMATION FOR SFO ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3336 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: US-08-977-553-5

Query Match 100.0%; Score 1029; DB 13; Length 3336;
Best Local Similarity 100.0%; Pred. No. 9, 2e-224;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Db 361 CGGCGATCACAATTATACCAAGACTTTTAAAGCAAAAGCAAGCACTGCCACAATT 420
Qy 421 ACAGTGAAGTATCTGCAAGTTTAGATGTAACAAAGCAATGTAATGACAAAGTCAA 480
Db 421 ACAGTGAAGTATCTGCAAGTTTAGATGTAACAAAGCAATGTAATGACAAAGTCAA 480
Qy 481 TGGATTACTAACAAGAGTTAAACAAAGATGCTATAAGTTAAGACATCGACAGCGCA 540
Db 481 TGGATTACTAACAAGAGTTAAACAAAGATGCTATAAGTTAAGACATCGACAGCGCA 540
Qy 541 GTGTTAACTGGAAGACGTACAGTTGAATATGATATCCAAATATACACGATTTCAA 600
Db 541 GTGTTAACTGGAAGACGTACAGTTGAATATGATATCCAAATATACACGATTTCAA 600
Qy 601 GATGAAAAAACCCTATTAAGTATATTTGCTAAGTGGGAAATATCATTTAAATCAG 660
Db 601 GATGAAAAAACCCTATTAAGTATATTTGCTAAGTGGGAAATATCATTTAAATCAG 660
Qy 661 CAAATTTATCAAGATGAATCAACACCAATTTGATATATACTGAAATCCAAATTTACA 720
Db 661 CAAATTTATCAAGATGAATCAACACCAATTTGATATATACTGAAATCCAAATTTACA 720
Qy 721 AGCAATCAACACATTTGAATTTATTACTTGAAGTCTTGATTTAACACATTTCTT 780
Db 721 AGCAATCAACACATTTGAATTTATTACTTGAAGTCTTGATTTAACACATTTCTT 780
Qy 781 CACAATTTATTAAGAGAGATGGAACCTTTGTAAGTGGAGGAGCCAAACACTACT 840
Db 781 CACAATTTATTAAGAGAGATGGAACCTTTGTAAGTGGAGGAGCCAAACACTACT 840
Qy 841 TCAGAAATTCCTCATTTATATATAGATCAATTTATCTATTAATGCCCCAAATTAAT 900
Db 841 TCAGAAATTCCTCATTTATATATAGATCAATTTATCTATTAATGCCCCAAATTAAT 900
Qy 901 GCGGATCTGGAATTTATCAATTTTATCAAAACAAATGATGATGACATGCCAGATGCG 960
Db 901 GCGGATCTGGAATTTATCAATTTTATCAAAACAAATGATGATGACATGCCAGATGCG 960
Qy 961 AACCAATTTGAATTTGATTCGAGTATTAATCAAAATGTAATTAATTAATTAATTCGA 1020
Db 961 AACCAATTTGAATTTGATTCGAGTATTAATCAAAATGTAATTAATTAATTAATTCGA 1020
Qy 1021 AAGAAGTGA 1029
Db 1021 AAGAAGTGA 1029

RESULT 9
US-08-977-866-7
: Sequence 7, Application US/08977866
: GENERAL INFORMATION:
: APPLICANT: Palmer, Leslie M.
: TITLE OF INVENTION: Novel ribb
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dechart Price & Rhoads
: STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
: City: Philadelphia
: STATE: PA
: COUNTRY: US
: ZIP: 19103
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/977,866
: FILING DATE:
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
```

FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Dickinson, Todd O  
 REGISTRATION NUMBER: 28,354  
 REFERENCE/DOCKET NUMBER: P50444-8  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 215-994-2252  
 TELEFAX: 215-994-2222  
 TELETYPE:  
 INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3336 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 US-08-977-866-7

Query Match 100.0%; Score 1029; DB 13; Length 3336;  
 Best Local Similarity 100.0%; Pred. No. 9,2e-224;  
 Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGATTTATGCGATTCACATTCGCAATATGCTACAGGTCAACAGAGTGTAAATCCACC 60  
 DB 1 ATGATTTATGCGATTCACATTCGCAATATGCTACAGGTCAACAGAGTGTAAATCCACC 60  
 QY 61 GTTGGCGCTGTGTAGTAAAGGTAGAGTGTGTGTATGCTGACAGCTTGAGAAAA 120  
 DB 61 GTTGGCGCTGTGTAGTAAAGGTAGAGTGTGTGTATGCTGACAGCTTGAGAAAA 120  
 QY 121 GGTGACACATGCGAGGTTCACAGACTTGTATGTCACACAAATATGCTGAAGTGGG 180  
 DB 121 GGTGACACATGCGAGGTTCACAGACTTGTATGTCACACAAATATGCTGAAGTGGG 180  
 QY 181 ACAGTTTATATACGTTAGAGCATGTAGCATTTTGTTCACACACACCCCTGTAAAC 240  
 DB 181 ACAGTTTATATACGTTAGAGCATGTAGCATTTTGTTCACACACACCCCTGTAAAC 240  
 QY 241 AAAATTTATGATTTAGATAGCAAAAGTATATACGCAACAAAGACAAATTCGTTAGAC 300  
 DB 241 AAAATTTATGATTTAGATAGCAAAAGTATATACGCAACAAAGACAAATTCGTTAGAC 300  
 QY 301 ACACATGCTGATGAGCGTTAGCGGCTCAGCGTATGAGGTGAATGCTGTATGAA 360  
 DB 301 ACACATGCTGATGAGCGTTAGCGGCTCAGCGTATGAGGTGAATGCTGTATGAA 360  
 QY 361 CGGCGATCACAATATACCAAGACTTTTAAAGCAAAAGCAAGACCTGCCACAAAT 420  
 DB 361 CGGCGATCACAATATACCAAGACTTTTAAAGCAAAAGCAAGACCTGCCACAAAT 420  
 QY 421 ACAGTGAAGATATCTGCAAGTGTAGATGTAACCAAGCAATGATATAGGACAAAGTCAA 480  
 DB 421 ACAGTGAAGATATCTGCAAGTGTAGATGTAACCAAGCAATGATATAGGACAAAGTCAA 480  
 QY 481 TGGATTACTAACAAGAGGTAAACAAGATGCTATAGTTAAGCATGACACGACGA 540  
 DB 481 TGGATTACTAACAAGAGGTAAACAAGATGCTATAGTTAAGCATGACACGACGA 540  
 QY 541 GTGTTAACTGAAGAGCTACAGTGAATAGATGATCCAAATATCTACAGTATTCAA 600  
 DB 541 GTGTTAACTGAAGAGCTACAGTGAATAGATGATCCAAATATCTACAGTATTCAA 600  
 QY 601 GATGCAAAAAACCCCTATAAAGTATATGCTAAGCTGGGAATATTCATTTAATCAG 660  
 DB 601 GATGCAAAAAACCCCTATAAAGTATATGCTAAGCTGGGAATATTCATTTAATCAG 660  
 QY 661 CAAATTTATCAAGATGAAATCAACACCAATTTGATATATACGAAATCCAAATTTAACA 720  
 DB 661 CAAATTTATCAAGATGAAATCAACACCAATTTGATATATACGAAATCCAAATTTAACA 720  
 QY 721 ACCAATCAACACATTTGAATTTACTGAACTCTTGATTAATTAACAACAAATTCCT 780  
 DB 721 ACCAATCAACACATTTGAATTTACTGAACTCTTGATTAATTAACAACAAATTCCT 780

QY 781 CACAATTTATATAAAGAGAGTGGACCTTGTAGTCAGAGCGAGTCCAACTACT 840  
 DB 781 CACAATTTATATAAAGAGAGTGGACCTTGTAGTCAGAGCGAGTCCAACTACT 840  
 QY 841 TCAGATTTCTCATCTATATATAGATGAATTTATCTCTATATAGCCCGAAATTAAT 900  
 DB 841 TCAGATTTCTCATCTATATATAGATGAATTTATCTCTATATAGCCCGAAATTAAT 900  
 QY 901 GCGGATCTGCAAAATATATCAACAAATGATGATGAGATGAGATACAGATGCG 960  
 DB 901 GCGGATCTGCAAAATATATCAATTTATCAACAAATGATGATGAGATGAGATGCG 960  
 QY 961 AACCAATTTGAAATTTGCTATCCGAGTTATTAATCAAAATGTTAAATTAATTACGA 1020  
 DB 961 AACCAATTTGAAATTTGCTATCCGAGTTATTAATCAAAATGTTAAATTAATTACGA 1020  
 QY 1021 AAGAGTGA 1029  
 DB 1021 AAGAGTGA 1029

RESULT 10  
 US-09-227-806-7  
 Sequence 7, Application US/09227806  
 GENERAL INFORMATION:  
 APPLICANT: Palmer, Leslie M.  
 APPLICANT: Pedon, Jason C.  
 APPLICANT: Warren, Richard L.  
 APPLICANT: Trainl, Christopher M.  
 APPLICANT: Wang, Min  
 APPLICANT: Jaworski, Deborah D.  
 APPLICANT: Mooney, Jeffrey  
 APPLICANT: Debouck, Christine  
 APPLICANT: Zhong, Yixi  
 APPLICANT: Black, Michael  
 TITLE OF INVENTION: TIDA  
 NUMBER OF SEQUENCES: 7  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Dechert, Price & Rhoads  
 STREET: 4000 Bell Atlantic Tower, 1717 Arch Ste  
 CITY: Philadelphia  
 STATE: PA  
 COUNTRY: USA  
 ZIP: 19103-2793  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSEQ for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/227,806  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/977,554  
 FILING DATE:  
 APPLICATION NUMBER: PCT/US97/02318  
 FILING DATE: 19-FEB-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Dickinson, O. Todd  
 REGISTRATION NUMBER: 28,354  
 REFERENCE/DOCKET NUMBER: P50444-07  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 215/994-2252  
 TELEFAX: 215/994-2222  
 TELETYPE:  
 INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3336 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear



US-09-227-806-7

Query Match 100.0%; Score 1029; DB 16; Length 3336;  
 Best Local Similarity 100.0%; Pred. No. 9.2e-224;  
 Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGATTATGCGATTCACTTGAATATGTGTACAGGTCAAGAGTGTATATCCACC 60
DB 1 ATGATTATGCGATTCACTTGAATATGTGTACAGGTCAAGAGTGTATATCCACC 60
QY 61 GTTGGCCCTGTGTAGTTAATGAAGGTAGCATTTGTGTATGTGTACACTTGAAGAA 120
DB 61 GTTGGCCCTGTGTAGTTAATGAAGGTAGCATTTGTGTATGTGTACACTTGAAGAA 120
QY 121 GGTGACAAAGCATGCGAGGTCAAGCACTTGTATGGACAAACAAATGCTGAAGGTGCG 180
DB 121 GGTGACAAAGCATGCGAGGTCAAGCACTTGTATGGACAAACAAATGCTGAAGGTGCG 180
QY 181 ACGATTATATTTACGTTAGAGCCATGTATGTCATTTTGTTCACACCACTCTGTATAC 240
DB 181 ACGATTATATTTACGTTAGAGCCATGTATGTCATTTTGTTCACACCACTCTGTATAC 240
QY 241 AAAATTATTTGATTGTAAGATAGCAAAAAGTACTATACGCAACAAAGCAATTCGTTAGAC 300
DB 241 AAAATTATTTGATTGTAAGATAGCAAAAAGTACTATACGCAACAAAGCAATTCGTTAGAC 300
QY 301 ACACATGCGATGAGAGGTATACGGGCTACGGTATGTAGAGTTGATGCTGATGATGAA 360
DB 301 ACACATGCGATGAGAGGTATACGGGCTACGGTATGTAGAGTTGATGCTGATGATGAA 360
QY 361 CGGCGATCACAATTTATACCAAGACTTTTAAAGCAAAAGCAAGCAACTGCGCAAAATT 420
DB 361 CGGCGATCACAATTTATACCAAGACTTTTAAAGCAAAAGCAAGCAACTGCGCAAAATT 420
QY 421 ACAATGAAATGTATGCAAGTTTATGATGTATTAACACAGCAATGATGCAAAAGTCAA 480
DB 421 ACAATGAAATGTATGCAAGTTTATGATGTATTAACACAGCAATGATGCAAAAGTCAA 480
QY 481 TGGATTACTAACAAGAGTTAAACAGATGTCTATTAAGTTAAGATTCGACAGACGCA 540
DB 481 TGGATTACTAACAAGAGTTAAACAGATGTCTATTAAGTTAAGATTCGACAGACGCA 540
QY 541 GTCTTAACTGGAAGACGTACAGTTGAATGTAGATGATCCAAATPACTACAGTATTCAA 600
DB 541 GTCTTAACTGGAAGACGTACAGTTGAATGTAGATGATCCAAATPACTACAGTATTCAA 600
QY 601 GATGGAAGAAACCTTATAAAGTATATTTGTCTAAGTCTGGGAATTTTCAATTTAATCAG 660
DB 601 GATGGAAGAAACCTTATAAAGTATATTTGTCTAAGTCTGGGAATTTTCAATTTAATCAG 660
QY 661 CAAATTTATCAAGTATCAACACCAATTTGATATATCTGAAATTCGCAATTTTAAACA 720
DB 661 CAAATTTATCAAGTATCAACACCAATTTGATATATCTGAAATTCGCAATTTTAAACA 720
QY 721 AGCAATCAACACATATTTGAATTTTACTTGAAGTCTGTGATTTTAAACAATTTCTT 780
DB 721 AGCAATCAACACATATTTGAATTTTACTTGAAGTCTGTGATTTTAAACAATTTCTT 780
QY 781 CACATTTATATTAAGAGAGAGCTTGAATTTGCTAGTCCAGGACGCTCCACCACTACT 840
DB 781 CACATTTATATTAAGAGAGAGCTTGAATTTGCTAGTCCAGGACGCTCCACCACTACT 840
QY 841 TCAGATTTTCATCTTATATATAGATGATTTTCTCTATTTATGCCCGCAATTTAAT 900
DB 841 TCAGATTTTCATCTTATATATAGATGATTTTCTCTATTTATGCCCGCAATTTAAT 900
QY 901 GCGGATCTGGAATTTATCAATTTTATCAAAACAATGATGTATGATACCAAGTCCG 960
DB 901 GCGGATCTGGAATTTATCAATTTTATCAAAACAATGATGTATGATACCAAGTCCG 960
QY 961 AACCAATTTGAATTTGTTCAATTCGAGTATTTAAATCAAAATGTTAAATTTACCTTACGA 1020
DB 961 AACCAATTTGAATTTGTTCAATTCGAGTATTTAAATCAAAATGTTAAATTTACCTTACGA 1020

```

```

DB 961 AACCAATTTGAATTTGTTCAATTCGAGTATTTAAATCAAAATGTTAAATTTACCTTACGA 1020
QY 1021 AAGAAGTGA 1029
DB 1021 AAGAAGTGA 1029

```

```

RESULT 11
US-09-369-700-7
: Sequence 7, Application US/09369700
: GENERAL INFORMATION:
: APPLICANT: Palmer, Leslie M.
: TITLE OF INVENTION: Novel ribg
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dechert Price & Rhoads
: STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
: CITY: Philadelphia
: STATE: PA
: COUNTRY: US
: ZIP: 19103
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/369,700
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/978,456
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Dickinson, Todd O
: REGISTRATION NUMBER: 28,354
: REFERENCE/DOCKET NUMBER: P50444-9
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 215-994-2252
: TELEFAX: 215-994-2222
: TELEX:
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3336 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
US-09-369-700-7

```

Query Match 100.0%; Score 1029; DB 17; Length 3336;  
 Best Local Similarity 100.0%; Pred. No. 9.2e-224;  
 Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGATTATGCGATTCACTTGAATATGTGTACAGGTCAAGAGTGTATATCCACC 60
DB 1 ATGATTATGCGATTCACTTGAATATGTGTACAGGTCAAGAGTGTATATCCACC 60
QY 61 GTTGGCCCTGTGTAGTTAATGAAGGTAGCATTTGTGTATGTGTACACTTGAAGAA 120
DB 61 GTTGGCCCTGTGTAGTTAATGAAGGTAGCATTTGTGTATGTGTACACTTGAAGAA 120
QY 121 GGTGACAAAGCATGCGAGGTCAAGCACTTGTATGGACAAACAAATGCTGAAGGTGCG 180
DB 121 GGTGACAAAGCATGCGAGGTCAAGCACTTGTATGGACAAACAAATGCTGAAGGTGCG 180
QY 181 ACGATTATATTTACGTTAGAGCCATGTATGTCATTTTGTTCACACCACTCTGTATAC 240
DB 181 ACGATTATATTTACGTTAGAGCCATGTATGTCATTTTGTTCACACCACTCTGTATAC 240
QY 241 AAAATTATTTGATTGTAAGATAGCAAAAAGTACTATACGCAACAAAGCAATTCGTTAGAC 300
DB 241 AAAATTATTTGATTGTAAGATAGCAAAAAGTACTATACGCAACAAAGCAATTCGTTAGAC 300

```

QY	301	ACACATGGTATGAGACGTTAACGGGCTACGAGTATGAGTTGAATGCGTGAATGATGA	360
Db	301	ACACATGGTATGAGACGTTAACGGGCTACGAGTATGAGTTGAATGCGTGAATGATGA	360
QY	361	CGGGCATCACAATTAATVACCAAGACTTTTTAAAGCAAAAGCAAGCAACTGCCACAATT	420
Db	361	CGGGCATCACAATTAATVACCAAGACTTTTTAAAGCAAAAGCAAGCAACTGCCACAATT	420
QY	421	ACAGTGAAGATATCTGCAAGTTTATGATGGTAAACAGCGATATATATGGACAAATGCA	480
Db	421	ACAGTGAAGATATCTGCAAGTTTATGATGGTAAACAGCGATATATATGGACAAATGCA	480
QY	481	TGGATTACTAACAAAGAGGTTAAACAAGATGTCATATAGTTAAGACATGACACGACGA	540
Db	481	TGGATTACTAACAAAGAGGTTAAACAAGATGTCATATAGTTAAGACATGACACGACGA	540
QY	541	GTTGTTAACTGGAAAGACGTACAGTTGAATTAAGATGATCCACAATATACTACAGTATTC	600
Db	541	GTTGTTAACTGGAAAGACGTACAGTTGAATTAAGATGATCCACAATATACTACAGTATTC	600
QY	601	GATGGAAGAAAACCTATTAAGATTAATTTGTCTAAGTCTGGGAATTTCTATTTAATCAG	660
Db	601	GATGGAAGAAAACCTATTAAGATTAATTTGTCTAAGTCTGGGAATTTCTATTTAATCAG	660
QY	661	CAAAATTTCAAGATGATCAACACCAATTTGGATATATCTGAAGAAATCCAAATTTAACA	720
Db	661	CAAAATTTCAAGATGATCAACACCAATTTGGATATATCTGAAGAAATCCAAATTTAACA	720
QY	721	AGCAATCAAAACACATATTTGAATTTATTTACTTGAAGTCTGTGTGATTTAAACAATCTT	780
Db	721	AGCAATCAAAACACATATTTGAATTTATTTACTTGAAGTCTGTGTGATTTAAACAATCTT	780
QY	781	CACAATTTATTAAGAGAGGTGGAACCTTGCTAGTGGACGACGTCAACCACTACT	840
Db	781	CACAATTTATTAAGAGAGGTGGAACCTTGCTAGTGGACGACGTCAACCACTACT	840
QY	841	TCAGAAATCTCCACTATTAATATATAGATGAATTAATCTCTAATTAATGCCGAATTAAT	900
Db	841	TCAGAAATCTCCACTATTAATATATAGATGAATTAATCTCTAATTAATGCCGAATTAAT	900
QY	901	GCGGATCTGGAATTAATCAATTTTATCAACAACAATGATGATGAAGATACCAGATGCG	960
Db	901	GCGGATCTGGAATTAATCAATTTTATCAACAACAATGATGATGAAGATACCAGATGCG	960
QY	961	AACCAATTTGAATTTGTCATCTCGAGATTAATTAATCAAAATGTTAAATTAACCTTACGA	1020
Db	961	AACCAATTTGAATTTGTCATCTCGAGATTAATTAATCAAAATGTTAAATTAACCTTACGA	1020
QY	1021	AAGAAGTGA 1029	
Db	1021	AAGAAGTGA 1029	
RESULT 12			
US-09-376-633-7			
: Sequence 7, Application US/09376633			
: GENERAL INFORMATION:			
: APPLICANT: Palmer, Leslie M.			
: TITLE OF INVENTION: Novel ribg			
: NUMBER OF SEQUENCES: 7			
: CORRESPONDENCE ADDRESS:			
: ADDRESSEE: Dechert Price & Rhoads			
: STREET: 4000 Bell Atlantic tower, 1717 Arch Stre			
: CITY: Philadelphia			
: STATE: PA			
: COUNTRY: US			
: ZIP: 19103			
: COMPUTER READABLE FORM:			
: MEDIUM TYPE: Diskette			
: COMPUTER: IBM Compatible			
: OPERATING SYSTEM: DOS			
: SOFTWARE: FastSeq for Windows Version 2.0			

```

: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/376.633
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/978.456
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Dickinson, Todd O
: REGISTRATION NUMBER: 28,354
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 215-994-2252
: TELEFAX: 215-994-2222
: TELEX:
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3336 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
:
: US-09-376-633-7

```

```

Db 661 CAAATTTATCAAGATGATCAACACCAATTTGGATATATCTGAAAATCCAAATTTAACA 720
Qy 721 AGCAATCAAAACACATATTTGAATTTATCTGAAAGCTCTGATTTTACAACAATTCCT 780
Db 721 AGCAATCAAAACACATATTTGAATTTATTTACTGAAAGCTCTGATTTTACAACAATTCCT 780
Qy 781 CACAATTTATTAATAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 781 CACAATTTATTAATAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Qy 841 TCACAATTTCTGCATCTATATATATATATATATATATATATATATATATATATATATAT 900
Db 841 TCACAATTTCTGCATCTATATATATATATATATATATATATATATATATATATATATAT 900
Qy 901 GCGGATCTGGAATTTATCAATTTTATCAACAATGATGATGATGATGATGATGATGATGATG 960
Db 901 GCGGATCTGGAATTTATCAATTTTATCAACAATGATGATGATGATGATGATGATGATGATG 960
Qy 961 AACCAATTTGAATTTGTCATTCGAGTTATTAATCAAAATGTTAAATTAATTAATTAATTA 1020
Db 961 AACCAATTTGAATTTGTCATTCGAGTTATTAATCAAAATGTTAAATTAATTAATTAATTA 1020
Qy 1021 AAGAAGTGA 1029
Db 1021 AAGAAGTGA 1029

```

RESULT 13  
US-09-492-479-7

```

; Sequence 7, Application US/09492479
; GENERAL INFORMATION:
; APPLICANT: Palmer, Leslie M.
; TITLE OF INVENTION: Novel ribb
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/492,479
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/977,866
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Todd Q
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: P50444-8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2252
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3336 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-09-492-479-7

```

Query Match 100.0%; Score 1029; DB 18; Length 3336;  
Best Local Similarity 100.0%; Pred. No. 9,2e-224;

	Matches 1029;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy 1	ATGATATATGCAATTCACATTCGAAATATGATACAGGTCAACAGAGTGTATATCCACC 60				
Db 1	ATGATATATGCAATTCACATTCGAAATATGATACAGGTCAACAGAGTGTATATCCACC 60				
Qy 61	GTTGGCGCTGTTGATTAATGAAGTAGATTTGTTGATTTGTTGTCACACTTGAGAAA 120				
Db 61	GTTGGCGCTGTTGATTAATGAAGTAGATTTGTTGATTTGTTGTCACACTTGAGAAA 120				
Qy 121	GCTGACAGCATGCGGAGCTCAAGCACTTGATATGSCAACAAATGCTGAGAGTCG 180				
Db 121	GCTGACAGCATGCGGAGCTCAAGCACTTGATATGSCAACAAATGCTGAGAGTCG 180				
Qy 181	ACGATTTATATACGTTGAGAGCATGTATGATTTGTTGTTGTTGTTGTTGTTGTTGTTG 240				
Db 181	ACGATTTATATACGTTGAGAGCATGTATGATTTGTTGTTGTTGTTGTTGTTGTTGTTG 240				
Qy 241	AAATTTATTTGATTTGTAAGATGCAAAAGTAGATATACGCAACAAAGACAAATTCGTTAGC 300				
Db 241	AAATTTATTTGATTTGTAAGATGCAAAAGTAGATATACGCAACAAAGACAAATTCGTTAGC 300				
Qy 301	ACACATGCTGATGAGACGCTTACGGGCTCAGCGTATGAGTTGATGCTGATGATGAA 360				
Db 301	ACACATGCTGATGAGACGCTTACGGGCTCAGCGTATGAGTTGATGCTGATGATGAA 360				
Qy 361	CGGCGATCACAATTTATACCAAGACTTTTAAAGCAAAAGCAACGACCTGCCAATTT 420				
Db 361	CGGCGATCACAATTTATACCAAGACTTTTAAAGCAAAAGCAACGACCTGCCAATTT 420				
Qy 421	ACAGTGAAGTATCTGCAAGTTGATGTTAAACAGGCAATGATATGACAAAGTCAA 480				
Db 421	ACAGTGAAGTATCTGCAAGTTGATGTTAAACAGGCAATGATATGACAAAGTCAA 480				
Qy 481	TGATTTACTACAAAGAGCTTAACACAGATGCTATGATTTAAGCATCGACAGACCA 540				
Db 481	TGATTTACTACAAAGAGCTTAACACAGATGCTATGATTTAAGCATCGACAGACCA 540				
Qy 541	GTTGTTAACTGGAAGACGATGATGATGATGATGATGATGATGATGATGATGATGATG 600				
Db 541	GTTGTTAACTGGAAGACGATGATGATGATGATGATGATGATGATGATGATGATGATG 600				
Qy 601	GATGAAAAAACCCTATTAAGATATATGCTAAGCTGGAATATTCATTTTAATCAG 660				
Db 601	GATGAAAAAACCCTATTAAGATATATGCTAAGCTGGAATATTCATTTTAATCAG 660				
Qy 661	CAATTTATCAAGATGATCAACACCAATTTGATATATCTGAAATTCGAAATTTAACA 720				
Db 661	CAATTTATCAAGATGATCAACACCAATTTGATATATCTGAAATTCGAAATTTAACA 720				
Qy 721	AGCAATCAAAACACATATTTGAATTTATTTACTGAAAGCTGCTGATTTAACAACATCTT 780				
Db 721	AGCAATCAAAACACATATTTGAATTTATTTACTGAAAGCTGCTGATTTAACAACATCTT 780				
Qy 781	CACAATTTATTAATAAG 840				
Db 781	CACAATTTATTAATAAG 840				
Qy 841	TCACAATTTCTGCATCTAT 900				
Db 841	TCACAATTTCTGCATCTAT 900				
Qy 901	GCGGATCTGGAATTTATCAATTTTATCAACAATGATGATGATGATGATGATGATGATGATG 960				
Db 901	GCGGATCTGGAATTTATCAATTTTATCAACAATGATGATGATGATGATGATGATGATGATG 960				
Qy 961	AACCAATTTGAATTTGTCATTCGAGTTATTAATCAAAATGTTAAATTAATTAATTAATTA 1020				
Db 961	AACCAATTTGAATTTGTCATTCGAGTTATTAATCAAAATGTTAAATTAATTAATTAATTA 1020				
Qy 1021	AAGAAGTGA 1029				
Db 1021	AAGAAGTGA 1029				

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	97.8%	Score	1006.4	DB	12	Length	1119
Best Local Similarity	99.1%	Pred.	No. 9.4e-219				
Matches	1023	Conservative	0	Mismatches	6	Indels	3
				Gaps			1
OY	1	ATGCATTATGCCATTCCAACTTGC	AAATATGTACAAAGGTTC	TAATCA	CCCC	60	
Db	87	ATGCATTATGCCATTCCAACTTGC	AAATATGTACAAAGGTTC	TAATCA	CCCC	146	
OY	61	GTTGCGCGTGTGTATGTAATGAG	CGAGATGTTGGTATG	GGGACAC	TTGAG	120	
Db	147	GTTGCGCGTGTGTATGTAATGAG	CGAGATGTTGGTATG	GGGACAC	TTGAG	206	
OY	121	GGTGCAACGATCGCGAGGTTCA	GCACCTTGATATG	GCACAACAA	AATCTG	180	
Db	207	GGTGCAACGATCGCGAGGTTCA	GCACCTTGATATG	GCACAACAA	AATCTG	266	
OY	181	ACGATTATATTAACGTGTAGAG	CATGTAGTCACTTTGGT	TCAACAC	CAACCCGT	240	
Db	267	ACGATTATATTAACGTGTAGAG	CATGTAGTCACTTTGGT	TCAACAC	CAACCCGT	326	
OY	241	AAATATTGATGTGAAGATAC	CAAAAGTAGTATAC	GCACAA	AAAAGAC	300	

Db	327	AAAATTATGTATGTAAAGTAGCAAAAAGTGTATACGAAACAAAAGACAAATTGCTTAGAC	386
QY	301	ACACATGCTGTATGACGTTCACGGGCTCACACGGTATGAGCTGAATGCTTATGTATGA	360
Db	387	ACACATGGTGTATGACGTTACGGGCTCACGATTTAGAGGTGAATGCTTATGTATGA	446
QY	361	CGGGCATCACAATTATACCAAGACTTTTTTAAAGCAAAAGCAAGCAACTGCGACAAAT	420
Db	447	CGGGCATCACAATTATACCAAGACTTTTTTAAAGCAAAAGCAAGCAACTGCGACAAAT	506
QY	421	ACATGGAAGATATGCAAGTTTATGATGGTAAACCAAGCATATGATATATGACAAAGTCAA	480
Db	507	ACAATGGAAGATATGCAAGTTTATGATGGTAAACCAAGCATATGATATATGACAAAGTCAA	566
QY	481	TGCAATTTACTAACAAAGAGGTTTAAACCAAGATGCTATTAAGTTAAGCATGACACAGACA	540
Db	567	TGCAATTTACTAACAAAGAGGTTTAAACCAAGATGCTATTAAGTTAAGCATGACACAGACA	626
QY	541	GTGTAACTGTGAAAGACGTACAGTTGATTAAGATGATCCACAATATATACACAGTATTCAA	600
Db	627	GTGTAACTGTGAAAGACGTACAGTTGATTAAGATGATCCACAATATATACACAGTATTCAA	686
QY	601	GATGGAACCAAAACCTATTAACAAAGTATTTGTCTTAAGTCTGGGAATTTTCATTTTATCAG	660
Db	687	GATGGAACCAAAACCTATTAACAAAGTATTTGTCTTAAGTCTGGGAATTTTCATTTTATCAG	746
QY	661	CAATTTATCATAGATGATCAACACCAAAATTTGGATATATACGAAATTCAAATTTTACA	720
Db	747	CAATTTATCATAGATGATCAACACCAAAATTTGGATATATACGAAATTCAAATTTTACA	806
QY	721	AGCAATCAACACATATTTGAATTTATTTACTTGAAGTCTTGTGATTTAACACAAATCTTT	780
Db	807	AGCAATCAACACATATTTGAATTTATTTACTTGAAGTCTTGTGATTTAACACAAATCTTT	866
QY	781	CACAAATTTATATAAAGAGGATTTGACACTTCTCTAGTGGAGCGAGTCCAACTACT	840
Db	867	CACAAATTTATATAAAGAGGATTTGACACTTCTCTAGTGGAGCGAGTCCAACTACT	926
QY	841	TCGAAATTCCTCCACTCT--ATTATATATGATGAATTTATTTCTCTATTAATGCCCGGAATTA	897
Db	927	TCGAAATTCCTCCACTCTATTTATTAATGATGAATTTATTTCTCTATTAATGCCCGGAATTA	986
QY	898	ATTGGCGGATCTGGAAATTTATCAATTTTATCAACAAATGATGATGATTAACACAT	957
Db	987	ATTGGCGGATCTGGAAATTTATCAATTTTATCAACAAATGATGATGATTAACACAT	1046
QY	958	GCGAACCAATTTGAAATTTGTTCAATTCGAGTATTAATCAAAATGTTAAATTAACCTTA	1017
Db	1047	GCGAACCAATTTGAAATTTGTTCAATTCGAGTATTAATCAAAATGTTAAATTAACCTTA	1106
QY	1018	CGAAAGCAAGTCA 1029	
Db	1107	CGAAAGCAAGTCA 1118	
RESULT 15			
US-09-611-529-2079			
; Sequence 2079, Application US/09611529			
; GENERAL INFORMATION:			
; APPLICANT: George H. Shimer, Jr.			
; APPLICANT: George H. Miller			
; APPLICANT: Roberta S. Hare			
; APPLICANT: Karen J. Shaw			
; TITLE OR INVENTION: Staphylococcus aureus Related Compositions and Methods			
; FILE REFERENCE: 1034/IC9650U1			
; CURRENT APPLICATION NUMBER: US/09/611,529			
; CURRENT FILING DATE: 2000-06-30			
; PRIOR APPLICATION NUMBER: US 09/417,811			
; PRIOR FILING DATE: 1999-10-14			
; PRIOR APPLICATION NUMBER: US 09/353,718			
; PRIOR FILING DATE: 1999-07-14			
; PRIOR APPLICATION NUMBER: US 09/266,557			



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2001, 11:48:14 : Search time 79.87 Seconds  
(without alignments)  
4404.654 Million cell updates/sec

Title: US-09-376-633-1

Perfect score: 1029

Sequence: 1 ATGATTTATGCGATTCACCT.....TAACTTTAGCAAGAAGTCA 1029

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 246011 seqs, 170942543 residues

492022

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_NA\_New.\*

1: /cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq.\*

2: /cgn2\_6/ptodata/1/pna/US06\_NEW\_COMB.seq.\*

3: /cgn2\_6/ptodata/1/pna/US08\_NEW\_COMB.seq.\*

4: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq.\*

5: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq.\*

6: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq.\*

7: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq.\*

8: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq.\*

9: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq.\*

10: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq.\*

11: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq.\*

12: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq.\*

13: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq.\*

14: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq.\*

15: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq.\*

16: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq.\*

17: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq.\*

18: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq.\*

19: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq.\*

20: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1000	97.2	7588	4 US-08-956-171C-142	Sequence 142, App
2	52.4	5.1	1250	5 US-09-739-449-1859	Sequence 1859, App
3	50.8	4.9	757	5 US-09-739-449-2286	Sequence 2286, App
4	50	4.9	70732	6 US-60-248-505-288	Sequence 288, App
5	49.8	4.8	1215	5 US-09-739-449-218	Sequence 218, App
6	49.8	4.8	506113	5 US-09-739-449-6666	Sequence 6666, App
7	47.8	4.6	518	5 US-09-833-381-1393	Sequence 1393, App
8	47.2	4.6	723	5 US-09-739-449-8012	Sequence 8012, App
9	47.2	4.6	70732	6 US-60-248-505-288	Sequence 288, App
10	47.2	4.6	70732	6 US-60-248-505-288	Sequence 288, App
11	47.2	4.6	70732	6 US-60-248-505-288	Sequence 288, App
12	47.2	4.6	70732	6 US-60-248-505-288	Sequence 288, App
13	46.2	4.5	776	5 US-09-739-449-1234	Sequence 1234, App
14	46.2	4.5	776	5 US-09-739-449-1234	Sequence 1234, App
15	44.2	4.3	902	5 US-09-739-449-1929	Sequence 1929, App
16	43.8	4.3	798	5 US-09-739-449-1929	Sequence 1929, App
17	43.6	4.2	126883	6 US-60-248-505-614	Sequence 614, App
18	43.2	4.2	543	5 US-09-739-449-8029	Sequence 8029, App
19	43.2	4.2	543	5 US-09-739-449-8029	Sequence 8029, App
20	43.2	4.2	1131	5 US-09-739-449-1953	Sequence 1953, App
21	42.6	4.1	780	5 US-09-739-449-451	Sequence 451, App
22	42.6	4.1	1006	5 US-09-739-449-621	Sequence 621, App
23	42.2	4.1	1040	5 US-09-739-449-2287	Sequence 2287, App
24	42.2	4.1	834	5 US-09-739-449-2256	Sequence 2256, App
25	41.4	4.0	43546	6 US-60-254-168-38	Sequence 38, App
26	41	4.0	1418	5 US-09-739-449-1613	Sequence 1613, App
27	40.8	4.0	756	5 US-09-739-449-870	Sequence 870, App

C 28	40.8	4.0	787	5 US-09-739-449-2501	Sequence 2501, App
C 29	40.8	4.0	1343	5 US-09-801-833-6840	Sequence 6840, App
C 30	40.6	3.9	780	5 US-09-739-449-451	Sequence 451, App
C 31	40.6	3.9	910	5 US-09-739-449-2366	Sequence 2366, App
C 32	40.6	3.9	9834	4 US-08-956-171C-337	Sequence 37, App
C 33	40.4	3.9	830	5 US-08-739-449-370	Sequence 370, App
C 34	40.2	3.9	3931	4 US-08-956-171C-342	Sequence 342, App
C 35	40.2	3.8	69513	6 US-60-248-823-16	Sequence 16, App
C 36	39.6	3.8	461	5 US-09-737-223-15443	Sequence 15443, App
C 37	39.6	3.8	325507	6 US-60-248-505-267	Sequence 267, App
C 38	39.4	3.8	786	5 US-09-739-449-2366	Sequence 2366, App
C 39	39.2	3.8	757	5 US-09-739-449-2500	Sequence 2500, App
C 40	39.2	3.8	711	5 US-09-739-449-77	Sequence 77, App
C 41	39	3.8	909	5 US-09-739-449-636	Sequence 636, App
C 42	39	3.8	5136	4 US-08-956-171C-332	Sequence 332, App
C 43	38.8	3.8	778	5 US-09-739-449-2435	Sequence 2435, App
C 44	38.8	3.8	98573	6 US-60-248-505-542	Sequence 542, App
C 45	38.6	3.8	4358	5 US-09-308-453-1	Sequence 1, App

#### ALIGNMENTS

RESULT 1  
US-08-956-171C-142/C  
Sequence 142, Application US/08956171C  
GENERAL INFORMATION:  
APPLICANT: Charles Kunsch  
GIL H. Choi  
Patrick S. Dillon  
Craig A. Rosen  
Steven C. Barash  
Michael R. Fannon  
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 5255  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/956,171C  
FILING DATE: 20-Oct-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/009,861  
FILING DATE: January 5, 1996  
APPLICATION NUMBER: 08/781,986  
FILING DATE: January 3, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoover, Kenley K.  
REGISTRATION NUMBER: 40,302  
REFERENCE/DOCKET NUMBER: PB248P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 610-5790  
TELEFAX: (301) 309-8439  
INFORMATION FOR SEQ. ID NO: 142:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7588 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 142:  
US-08-956-171C-142















**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2001, 10:35:04 ; Search time 1372.91 Seconds  
(Without alignments)  
11053.563 Million cell updates/sec

Title: US-09-376-633-1  
Sequence: 1 ATGGATTCGATTCGACT.....TAACCTTACGAAGACGTGA 1029

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1283235 seqs, 7373929652 residues

Total number of hits satisfying chosen parameters: 2566470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_ba1: \*  
2: gb\_ba2: \*  
3: gb\_ba3: \*  
4: gb\_in1: \*  
5: gb\_in2: \*  
6: gb\_in3: \*  
7: gb\_om: \*  
8: gb\_ov: \*  
9: gb\_pat1: \*  
10: gb\_pat2: \*  
11: gb\_ph: \*  
12: gb\_p11: \*  
13: gb\_p12: \*  
14: gb\_p13: \*  
15: gb\_p14: \*  
16: em\_ba1: \*  
17: em\_ba2: \*  
18: em\_fun: \*  
19: em\_htgo\_hum: \*  
20: em\_htgo\_inv: \*  
21: em\_htgo\_rod: \*  
22: em\_htg\_hum1: \*  
23: em\_htg\_hum2: \*  
24: em\_htg\_hum3: \*  
25: em\_htg\_hum4: \*  
26: em\_htg\_hum5: \*  
27: em\_htg\_hum6: \*  
28: em\_htg\_hum7: \*  
29: em\_htg\_hum8: \*  
30: em\_htg\_inv1: \*  
31: em\_htg\_inv2: \*  
32: em\_htg\_other: \*  
33: em\_htg\_rod: \*  
34: em\_hum1: \*  
35: em\_hum2: \*  
36: em\_hum3: \*  
37: em\_hum4: \*  
38: em\_hum5: \*  
39: em\_hum6: \*  
40: em\_hum7: \*  
41: em\_in: \*  
42: em\_om: \*  
43: em\_ot: \*

44: em\_ov: \*  
45: em\_pat: \*  
46: em\_ph: \*  
47: em\_pl: \*  
48: em\_ro: \*  
49: em\_sts: \*  
50: em\_sy: \*  
51: em\_un: \*  
52: em\_vi: \*  
53: gb\_sts1: \*  
54: gb\_sts2: \*  
55: gb\_sts3: \*  
56: gb\_sy: \*  
57: gb\_un: \*  
58: gb\_v11: \*  
59: gb\_v12: \*  
60: gb\_hvg1: \*  
61: gb\_hvg2: \*  
62: gb\_hvg3: \*  
63: gb\_hvg4: \*  
64: gb\_hvg5: \*  
65: gb\_hvg6: \*  
66: gb\_hvg7: \*  
67: gb\_hvg8: \*  
68: gb\_hvg9: \*  
69: gb\_hvg10: \*  
70: gb\_hvg11: \*  
71: gb\_hvg12: \*  
72: gb\_hvg13: \*  
73: gb\_hvg14: \*  
74: gb\_hvg15: \*  
75: gb\_hvg16: \*  
76: gb\_hvg17: \*  
77: gb\_hvg18: \*  
78: gb\_hvg19: \*  
79: gb\_hvg20: \*  
80: gb\_hvg21: \*  
81: gb\_hvg22: \*  
82: gb\_hvg23: \*  
83: gb\_hvg24: \*  
84: gb\_hvg25: \*  
85: gb\_pr1: \*  
86: gb\_pr2: \*  
87: gb\_pr3: \*  
88: gb\_pr4: \*  
89: gb\_pr5: \*  
90: gb\_pr6: \*  
91: gb\_pr7: \*  
92: gb\_pr8: \*  
93: gb\_pr9: \*  
94: gb\_rol: \*  
95: gb\_rol2: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1029	100.0	3336	9 AR069579	AR069579 Sequence
2	516.8	50.2	3227	2 AF269646	AF269646 Staphyloc
3	516.8	50.2	3618	2 AF269345	AF269345 Staphyloc
4	340.6	33.1	2902	2 AF270130	AF270130 Staphyloc
5	155	15.1	3721	2 AF269926	AF269926 Staphyloc
6	148	14.4	300550	2 AP001512	AP001512 Bacillus
7	121.2	11.8	293181	2 AP001119	AP001119 Buchnera
8	120.4	11.7	6006	2 BSRIB	X51510 B. subtilis
9	120.4	11.7	28206	2 BACDIA	L09228 Bacillus su
10	120.4	11.7	218470	2 BSDB0013	Z99116 Bacillus su
11	116.4	11.3	14364	1 AE000675	AE000675 Aquifex a

```
12 112.2 10.9 4286 2 BARIBGENS X95955 B.amyloliqu
13 111 10.8 175936 84 SPNE01908 AL449930 Streptoco
14 105.8 10.3 11167 1 AE006333 Lactococc AE006333
15 95.2 9.3 18073 3 U32775 Haemophilus U32775
16 95.4 8.9 11695 1 AE004298 Vibrio ch AE004298
17 85.8 8.3 10166 1 AE006112 Pasteurel AE006112
18 83.4 8.1 12811 1 AE001819 Thermotog AE001819
19 79 7.7 5451 3 TNE9832 Thermotog AE009832
20 78.6 7.6 2949 2 ECNUSB X64395 E.coli gene
21 78.6 7.6 11692 1 AE000148 AE000148 Escherich
22 78.6 7.6 139818 2 ECUR2664 D82664 Escherichia
23 77.4 7.5 105795 1 SYCCPBC U27202 Actinobacti
24 76.2 7.4 4312 2 AP027202 U27202 Actinobacti
25 75.4 7.3 2759 1 AE002857 Shigella AE002857
26 75.4 7.3 10323 1 AE005220 AE005220 Escherich
27 71.2 6.9 269223 10 AX067466 AX067466 Sequence
28 67.8 6.6 11456 1 AE002377 Chlamydia AE002377
29 67.6 6.6 18759 1 AE002257 Chlamydia AE002257
30 67.6 6.6 20386 1 AE001667 Chlamydia AE001667
31 67.6 6.6 32585 2 AP002548 Chlamydia AP002548
32 66 6.4 11544 1 AE001343 Chlamydia AE001343
33 65.4 6.4 14376 1 AE002531 Neisseria AE002531
34 65.4 6.4 34980 9 AX044034 AX044034 Sequence
35 62.4 6.1 349061 3 NMA223491 A1162753 Neisseria
36 60.4 5.9 2307 2 BHE132928 A1132928 Bartonell
37 60.4 5.9 85779 14 SCE011856 A1011856 Sacccharom
38 60.2 5.9 86827 6 PEMALP5 AL034556 Plasmodiu
39 59.8 5.8 964 54 CNS07E8R AL441457 T7 end of
40 59.6 5.8 11404 1 AE003934 AE003934 Xylella f
41 58 5.6 60604 68 AC023466 AC023466 Homo sapi
42 56.4 5.5 150236 91 HS101G11 AL021877 Human DNA
43 56.4 5.5 205429 60 AC005506 AC005506 Plasmodiu
44 55.6 5.4 7218 10 166494 Sequence 14
45 55.6 5.4 162160 63 AC013817 AC013817 Homo sapi
```

## ALIGNMENTS

```
RESULT 1
AR069579 AR069579 3336 bp DNA PAT 18-FEB-2000
LOCUS Definition Sequence 7 from patent US 5891672.
ACCESSION AR069579
VERSION AR069579.1 GI:7220467
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 3336)
Wang,M., Ward,J.M., Warren,R.,Lloyd, Nicholas,R.,Oakley,
Palmer,L.,Marle, Pratt,J.M., Knowles,D.,Justincharles,
Lometto,M.,Arthur, Mooney,J., Black,M.,Terence,
Burnham,M.,KarlRussell, Debouck,C., Fedon,J.,Craig, Hodgeson,J.,Edward,
Jaworski,D.,Dee, Reichard,R.,Winfield, Rosenberg,M., Trainl,C.,Michael
and Zhong,Y.Xi.
TITLE Polynucleotides encoding GTP cyclohydrolase II (riba)
JOURNAL Patent: US 5891672-A 7 06-APR-1999;
FEATURES Location/Qualifiers
source 1..3336
BASE COUNT 1202 a 459 c 678 g 997 t
ORIGIN
```

Query Match 100.0%: Score 1029; DB 9; Length 3336;  
Best Local Similarity 100.0%: Pred. No. 6.3e-179;  
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGATATGCGATTCACACTTGCATAATATGATACAAAGTCAAGAGTGTATATCCACCC 60
DB 1 ATGATATGCGATTCACACTTGCATAATATGATACAAAGTCAAGAGTGTATATCCACCC 60
```

```
QY 61 GTTGGCGCTGTGTACTATATAGAGTAGAGTGTGTGTTGTCGCACCTTGAGAAA 120
DB 61 GTTGGCGCTGTGTACTATATAGAGTAGAGTGTGTGTTGTCGCACCTTGAGAAA 120
QY 121 GGTGACAAACATGCGGAGGTTCAGACCTGATATGGCAACAAATCTGAAGTGGC 180
DB 121 GGTGACAAACATGCGGAGGTTCAGACCTGATATGGCAACAAATCTGAAGTGGC 180
QY 181 ACAGATTTATATACGTTAGAGCAGTATAGTCAATTTGGTTCACACACCTGTGTAA 240
DB 181 ACAGATTTATATACGTTAGAGCAGTATAGTCAATTTGGTTCACACACCTGTGTAA 240
QY 241 AAAATTTATGATTTAGATATACGAAAAGTATACGACAAAGACAAATTCGTTAGC 300
DB 241 AAAATTTATGATTTAGATATACGAAAAGTATACGACAAAGACAAATTCGTTAGC 300
QY 301 ACACATGATGATGAGAGCTTACGGGCTCAGGATTTAGGTTGAATGGCTGATATGA 360
DB 301 ACACATGATGATGAGAGCTTACGGGCTCAGGATTTAGGTTGAATGGCTGATATGA 360
QY 361 CCGGCATCAATTTATACCAAGCTTTTAAAGCAAAAGCAAGCAACTGCCAATTT 420
DB 361 CCGGCATCAATTTATACCAAGCTTTTAAAGCAAAAGCAAGCAACTGCCAATTT 420
QY 421 ACAGTGAAGATATCTGCAAGTTAGATGTTAAACAAGCAATGATATGCAAGTCAA 480
DB 421 ACAGTGAAGATATCTGCAAGTTAGATGTTAAACAAGCAATGATATGCAAGTCAA 480
QY 481 TGGATTTACATCAAAAGAGTTTAAACAGATGCTATTAAGTTAAGCATGACACGCA 540
DB 481 TGGATTTACATCAAAAGAGTTTAAACAGATGCTATTAAGTTAAGCATGACACGCA 540
QY 541 GTGTTTAACTGGAAGAGTACAGTTGAATTTAGATTCACAAATATACGATATTCAA 600
DB 541 GTGTTTAACTGGAAGAGTACAGTTGAATTTAGATTCACAAATATACGATATTCAA 600
QY 601 GATGAAAAAACCCTRTAAAGTATATTTGCTAAGTCTGGCAATTTATTTATTCG 660
DB 601 GATGAAAAAACCCTRTAAAGTATATTTGCTAAGTCTGGCAATTTATTTATTCG 660
QY 661 CAAATTTATCAAGATGATCAACACCAATTTGATATATCTGAAATTTATTTAACA 720
DB 661 CAAATTTATCAAGATGATCAACACCAATTTGATATATCTGAAATTTATTTAACA 720
QY 721 AGCAATCAAAACATATTTGAAATTTATTTGAAAGTCTTGATTTTAAACAATTC 780
DB 721 AGCAATCAAAACATATTTGAAATTTATTTGAAAGTCTTGATTTTAAACAATTC 780
QY 781 CACATTTATATATAAGAGAGAGTGGAACTTTGAGTCTAGTGGAGGAGGTCACACT 840
DB 781 CACATTTATATATAAGAGAGAGTGGAACTTTGAGTCTAGTGGAGGAGGTCACACT 840
QY 841 TCAGATTTCTCATCTATATATATAGATGATTTATTTCTATATATGCCGGAATTA 900
DB 841 TCAGATTTCTCATCTATATATATAGATGATTTATTTCTATATATGCCGGAATTA 900
QY 901 GCGGATCTGGAATTTATCAATTTTATCAACAATGATGATGATGATGATGATGAT 960
DB 901 GCGGATCTGGAATTTATCAATTTTATCAACAATGATGATGATGATGATGATGAT 960
QY 961 AACCAATTTGAATTTGTTATTCGAGATTTAAATCAAAAGTAAATTAATTAATTA 1020
DB 961 AACCAATTTGAATTTGTTATTCGAGATTTAAATCAAAAGTAAATTAATTAATTA 1020
QY 1021 AAGAGTGA 1029
DB 1021 AAGAGTGA 1029
```

```
RESULT 2
AF269646/c AF269646 3227 bp DNA BCT 01-AUG-2000
LOCUS AF269646/c
DEFINITION Staphylococcus epidermidis strain SRI clone step.1015d10 genomic
```



ACCESSION	sequence.
VERSION	AF269646
KEYWORDS	AF269646.1 GI:9623542
SOURCE	
ORGANISM	Staphylococcus epidermidis. Staphylococcus epidermidis Bacteria: Firmicutes: Bacillus/Clostridium group; Bacillus/Staphylococcus group; Staphylococcus. 1 (bases 1 to 3227)
REFERENCE	
AUTHORS	Kimmely,W.J., Taylor,J.David, Nelsen,A.J., Godlevski,M.M.,

**TITLE** Transposon-mediated sequencing of the *Staphylococcus epidermidis* genome

JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 3227)
AUTHORS	Taylor, J. David, Kimmerly, W. J., Nelsen, A. J., Godlewski, M. M.,

**FEATURES**

<b>TITLE</b>	Direct Submission
<b>JOURNAL</b>	Submitted (22-MAY-2000) Departments of Genomic Sciences and Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore Drive, Research Triangle Park, North Carolina 27709-3398, USA
<b>Location/Qualifiers</b>	

source 1. .3227

BASE COUNT	980 a	599 c	454 g	1194 t
ORIGIN				

Query Match	50.2%	Score 516.8	DB 2	Length 3227
Best Local Similarity	69.5%	Pred. No. 2.9e+85		
Matches 717	Conservative 0	Mismatches 312	Indels 3	Gaps 1

OY	1	ATATATTGGCATTTCACTTGGCAATATGGTACAAAGTCACAGGCTTAATCCACC	60
Db	1918	ATGGATGATGCTATTTCACACTACGCAAAAATGGTAATGACAAACAGGCTTAATCCACCA	1859
OY	61	GTTGGCGCTGTTTGTAGTAAATGAAGGTAGGATGTTGGTATTTGGTCACACTTGGAGAAA	120
Db	1858	GTAGAGATCCGTTGTTGTTAAAAACGGTARGANTGTAGGTTTAGTGTCACATTAAAAAAG	1799
OY	121	GGTGACAAAGCATGGCGAGGTTTCAGCCACTTGATATGSCAACAAATAGCTGAGGTGGC	180
Db	1798	GGAATATAAACATGCGCGAAGTACAAAGCTTTGTAAGGACAGGCTTTAAATACCCAAAGGTGCT	1739
OY	181	ACGATTTTATTACGTTAGAGCCATGTATGTCATTTTGGTTCACACCAACCCGTGTATAC	240
Db	1738	ACCATATACGTTTCATTATGAACCTTGGCACACACCATGTTGCACACCACCTTGTTGTCAT	1679
OY	241	AAATATTATGATTGTAGATAGCAAAAGTAGTATAGCACAACAAAGACAAATTCGTTAGAC	300
Db	1678	AAATTCATTGAAGCGGCGCATATCTAACGTCAATCTATGCTTTAAAGATACACTTATAGTA	1619
OY	301	ACACATGCGTATGAGAGCTTACCGGGCTCACGCTATTTAGGTTGGATGCGTTAGATATGAA	360
Db	1618	AGTAAAGGTGACGABATTTCTGAGAGAAAGCTGTAATAGAGTTGAAATTTCAATATATATGAA	1559
OY	361	CGGCGATCACAATTAATATACCAAGACTTTTAAAAAGCAAAAGCAAAAGCAACATGCGACAAATT	420
Db	1558	AATCCAGCTGCATTAATACCGTGACCTTTTCTACTGTTAAAAAGAAAGAAAGATTCCAAAGTA	1499
OY	421	ACAGTAAAGATATCGCAAGTTTATGATGTTAAACAAGCAAGGAATGATTAATGACAAAGTCAA	480
Db	1498	ACTGTAAAGGTCATCACTAGTCTGATGTGTAATACAGCAACAGACTTTTAATGCAAAAGTAAG	1439

[illegible]

RESULT	3
AF269345/c	
LOCUS	3618 bp DNA
DEFINITION	Staphylococcus epidermidis strain SRI clone step.1002d07 genomic sequence.
VERSION	AF269345
KEYWORDS	AF269345.1 GI:9664666
SOURCE	Staphylococcus epidermidis.
ORGANISM	Staphylococcus epidermidis. Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Staphylococcus.
REFERENCE	1 (bases 1 to 3618) Kimmerly,W.J., Taylor,J.David, Nelsen,A.J., Godlevski,M.M., Riddio,M.A., Nelson,F.J., Rivers,P.R., Torruebla-Miller,I., Listenbee,S., Ashanti,C., Altschuller,G., Mamola,L., Shepherd,N.S., Fuchs,R., Fleming,T., Guan,X., Du,L., Calh,D.H., Miller,G.S. and Furdon,P.J.
TITLE	Transposon-mediated sequencing of the Staphylococcus epidermidis genome
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 3618) Taylor,J.David, Kimmerly,W.J., Nelsen,A.J., Godlevski,M.M., Riddio,M.A., Nelson,F.J., Rivers,P.R., Torruebla-Miller,I., Listenbee,S., Ashanti,C., Altschuller,G., Mamola,L., Shepherd,N.S., Fuchs,R., Fleming,T., Guan,X., Du,L., Calh,D.H., Miller,G.S. and Furdon,P.J.
JOURNAL	Direct Submission
FEATURES	Submitted (19-MAY-2000) Departments of Genomic Sciences and Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore Drive, Research Triangle Park, North Carolina 27709-3398, USA Location/Qualifiers

```

source      1. .3618
            /organism="Staphylococcus epidermidis"
            /strain="SRI"
            /db_xref="taxon:1282"
            /clone="step.1002407"
BASE COUNT  1099 a      709 c      518 g      1292 t
ORIGIN

```

Query Match	50.2%;	Score 516.8;	DB 2;	Length 3618;
Best Local Similarity	69.5%;	Pred. No. 2.8e-85;		
Matches 717; Conservative	0;	Mismatches 312;	Indels 3;	Gaps 1;

QY	1	ATGATTTATGGGATTCACAACTTGGAAATATGGTTCAGAGTCAACAGAGTTTATCATACC	60
Db	2701	ATGATATGATGCTATTTCACACTAGCAAAAATGGTAAATGGCAACACAGGTTTATCATCCCA	26424
QY	61	GTTCGGCGCTGTTGTAGTTAAATGAAGGTAGAGATTGTTGGTATTGGTGCACACTTGGAGAAA	120
Db	2641	GTAGGATCCGCTGTGTTTAAAAAACGGTAGAGTTGTAGGTTTATAGTGCACATTTAAAAAAG	25824
QY	121	GGTACACACGATGGGAGGTTTCAAGCACTTGATATGGCAACAAATATGCTGAAAGTGGC	180
Db	2581	GGAAATTAACATGCGCGGAATGCAAGCTATTGGAATGCGACGTTTAAATACCAACAGTCTCT	25322
QY	181	ACGATTTATATATACGTTTAGAGCCCATGTAGTCAATTTTGGTTCACACCAACCCGTGTATAC	240
Db	2521	ACCATATACGTTTTCATTATAGAACCTTGACACACCAATGTTTCAACACACACTGTGTGCAT	24622
QY	241	AAATTTATTCGTTTCTAAGATAGCAAAAAGTATGATACGCAACAAAAGACAAATTCGTTAGAC	300
Db	2461	AAAATTCATTAAGCCGGCAATATATTAAGGTCATCATGCTTTAAAGATATCACTTTTACTA	24020
QY	301	ACACATGCGTATGAGACGTTTACCGGCTCACAGGATATGAGGTTGAATGGGTGATGATGAA	360
Db	2401	AGTAAGGCTGACAGATTCGTGACAGAAAGCGGATATAGAGCTTGATTTCAATATATATCA	23424
QY	361	CGGGCATTCACATTTATACCAAGACTTTTTTAAAGCAAAAAGCAACCAACTGCCACAAAT	420
Db	2341	AATGACACTGCATTTATACCGTGCATTTTTTACTGCTTAAAGAAAGCAAGAGTTCCAGAACTA	22822
QY	421	ACAGTGAAGATATCTCGCAATTTAGATGGTGAACCAAGCAAGATATATGACAAAGTCAA	480
Db	2281	ACTGTAAAGGCTCATCTACTACTAGTATGTTAAACAAAGCAACAGACTTTTAAATGAAGTAA	22222
QY	481	TGGATTACTAACAAGAGGTTTAAACAAGATGCTATAGTTAAGATCATGCACACAGCA	540
Db	2221	TGCGATTACAACAACAAAGAAAGTTAAAGAAAGTATTATCATTTAAGCATAGACATGATATCA	21622
QY	541	GTTTAACTCGGAAGACCTACAGTTGAATATAGATATCCCAATATATCATACACGTATTGAA	600
Db	2161	GTATTTACTGGGGGTAAACCATTTGAAGCAAGCAACATCCATTTATATCAACCAAGGGTTCT	21022
QY	601	GATGAGAAAAAACCCTATTAAGAAAGTAAATATGTTCTAAGTCTGGCAATATTCATTTTAAATAG	660
Db	2101	GATGGAAGAAGATCGGATTCGAGTTATTTCTTTCTTAAGAAAGTCACTCGATTTTATATCAA	20422
QY	661	CAAAATTTATCAGATGATCATAACCAAAATTTGGATATATAGTGAAGAAATCCAAATTTAACA	720
Db	2041	CAAAATTTAAAGATATCTGCATCGAGATATGATATTCACCTGAAAAAATGAAGAAATTTAAA	19822
QY	721	AGCAATCAACAACATATTTGAATTTATTTCTTGTAAGCTCTTGATTTTAAACAACATTTCT	780
Db	1981	ACAAATTAAGATTTTATTTAAATATATATATATAGTAAATTTGATATACACGCAATATATTA	19222
QY	781	CACAATTTATATTAAGAGGAGTTTGAAGCTTTGCTAGTCGAGGCGAGTCCAAACACTACT	840
Db	1921	CAGACCTTATATCAAGAGAGTATTTGGGAAAGCTGTAGTCGAGGCGAGGCCAAATATTACA	18622
QY	841	TCGAAATTTCTCCATCT---ATATATATAGATGAATTTATTTCTATATATATGCCCCGAAATTA	897
Db	1861	TCCTCAATTTCTCAATCCAAACCTCTTAATATGAATCTATTTTATATATAGCCCCGAAATTA	18022

QY 898 ATGGCGCGATCTCGAATAATTATCAATTTTATCAACAAGATGATGGATTTGGATACCCAGAT 957

Db 1801 ATTTGGTGGTCTGGCAACATCAATTTTATTAAGACTGCGAGGTCATTGGATTGGCCCTGAA 1742

QY 958 GCGACCAATTAAATGTCATTCCGAGTATTAAATCAAAATGTTAAATTAACCTTA 1017

Db 1741 GCAACTCAATTTGAATGTTGATTTCCAGATTAAATCAAAATTTAAATTAAGAAATTA 1682

QY 1018 CGAAGAGATGA 1029

Db 1681 CGAAGAGATGA 1670

Qy	1018	CGAAGAGACTGA	1029
Db	1681	CGAAGAGACTGA	1670

RESULT	4
LOCUS	AF270130
DEFINITION	AF270130 2902 bp DNA BCT 01-AUG-2000 Staphylococcus epidermidis strain SFI clone step.1051a10 genomic sequence.
ACCESSION	AF270130
VERSION	AF270130.1 GI:9624037
KEYWORDS	
SOURCE	Staphylococcus epidermidis.
ORGANISM	Staphylococcus epidermidis

REFERENCE  
AUTHORS  
1 (bases 1 to 2902)  
Kimmerly, W.J., Taylor, J. David, Nelsen, A.J., Godlevski, M.M.,

TITLE	Transposon-mediated sequencing of the <i>Staphylococcus epidermidis</i> genome
-------	--

JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 2902)
AUTHORS	Taylor, J. David, Kimmerly, W. J., Nelsen, A. J., Godlevski, M. M.,

**TITLE** Direct Submission  
**JOURNAL** Submitted (22-MAY-2000) Departments of Genomic Sciences and  
 Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore  
 Drive, Research Triangle Park, North Carolina 27709-3398, USA  
**FEATURES** Location/Qualifiers

```

/organism="Staphylococcus epidermidis"
/strain="SK1"
/db_xref="taxon:1282"
/clone="step_1051a10"
BASE COUNT      1009 a      382 c      507 g      1004 t
ORIGIN

```

Query Match	33.1%	Pred. 340.6	DB 2	Length 2902
Best Local Similarity	69.0%	Pred. No. 4.8e-53		
Matches 466; Conservative	0	Mismatches 209	Indels 0	Gaps 0

QY	1	ATGATTTATGCGATTCACACTTGGCAATATNGTACAAAGGTCAACACAGGTGTATATCCACC	60
Db	2220	ATGATGATATCTATTCACACACGAAAAATGTAAATGCAACACAGGTGTATATCCACA	2279
QY	61	GTTGGCCGTGTGTAGTATATGAAGGTAGAGATGTGTGATATGGGACACACTTGGAGAA	120
Db	2280	GTAGGATCCGTTGTTGTTAAAAACGGTAGGATTTGTAGGTTTAGGTGCACATTTAAAAAAG	2338
QY	121	GGTGACACAGCATCGGAGGTTTCAAGCACTTGATATGGCAACAACAAATGCTGGAAGTGGC	180
Db	2340	GGAGATTAACATGGCCGAAGTACAAAGTATTTGAATGCGAGTTTAAATCCCAAGCTGCT	2399
QY	181	ACGATTTATATACGTTTAGAGCCATGACTGATTTTGGTTCAACACACCCCTGTGTTAAC	240
Db	2400	ACCATATACGTTTCATATAGAACCTTTCACACACACATAGTTTCAACACACACTGTGTGCAT	2458

QY	241	AAATTTTGTATGTAAAGTAAAGCAAAAGTAAGTACCAACAAAGACAATTCGTTAGAC	300
Db	2460	AAAATCATTTGAAGGGGCATATCTAAAGTTCATCTATGCTGTTAAAGATCTACTTTAGTA	2519
QY	301	ACACATGSGTATGTAGACGTTACGGGCTCAACGGTATTGAGTTGAATGGCTGTATGATGATGA	360
Db	2520	AGTAAGGAGTCGACGATTTGTGAGAGAAGAGCTGGTATAGAGTTGGAATTTCAATATATATGA	2579
QY	361	CGGAGCATCACATATATATCCAAAGACTTTTTTAAAGCAAAAGCAACACACTCGCCAAATT	420
Db	2580	AATCAGACTGCATTAATATCCGTGACTTTTTTACTGTCTAAAGAAAGCAAGAGTTCGCAAGTA	2639
QY	421	ACAGTGAAGATATCTGCCAACTTTAGATGTGAACAAGCAAGCAATGATATAGACAAAGTGA	480
Db	2640	ACTGTAAAGGTCATCTACTACTAGATGGAACAAAGCAACAGACTTTTAATGAAGATAG	2699
QY	481	TGATTTACTAACAAGAGGTTAAACAAGATGTCTATATAGTTAAGCATCGACAGACGA	540
Db	2700	TGGATTAACAACAAGAAAGTTAAAGAAAGTATTTTATCAATTTAAGCAATGACATGATGATCA	2759
QY	541	GTGTAACTGGAAGACTTACGATTGAATTAGATGATCCCAATATATACACGATTTCA	600
Db	2760	GTTATTACTGGGCTAGAACCATATGAAGCATCATCTTTATATACCAACGAGGTTCT	2819
QY	601	GATGGAAGAAAAACCTATATAAAGTAAATATTTGTCTAAGTCTGGGAATATTCATTTTAATCG	660
Db	2820	GATGGAAGACATCCGATTCAGATTTATCTTTCTTAAGAAAGTCACACTCGATTTTAATCAA	2879
QY	661	CAATTTATCAAGT	675
Db	2880	CAATATTTTAAAGT	2894

RESULT	5	
LOCUS	AF269926/c	
DEFINITION	AF269926 3721 bp DNA BCT 01-AUG-2000	
ACCESSION	Staphylococcus epidermidis strain SRI clone step.1032d02 genomic sequence.	
VERSION	AF269926	
KEYWORDS	AF269926.1 GI:9623827	
SOURCE		
ORGANISM	Staphylococcus epidermidis. Staphylococcus epidermidis. Bacteria; Firmicutes; Bacillus/Clostridium group; Staphylococcus.	
REFERENCE	1 (bases 1 to 3721) Kimmerly W.J., Taylor J.David, Nelsen A.J., Godlewski, M.M., Rudino, M.A., Nelson, F.J., Rivers, P.R., Torruella-Miller, I., Listebdee, S., Ashanti, C., Altschuller, G., Mamo, L., Shepherd, N.S., Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D.H., Miller, G.S. and Furdon, P. J.	
AUTHORS	Transposon-mediated sequencing of the Staphylococcus epidermidis genome	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 3721)	
AUTHORS	Taylor, J.David, Kimmerly, W.J., Nelsen, A.J., Godlewski, M.M., Rudino, M.A., Nelson, F.J., Rivers, P.R., Torruella-Miller, I., Listebdee, S., Ashanti, C., Altschuller, G., Mamo, L., Shepherd, N.S., Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D.H., Miller, G.S. and Furdon, P. J.	
TITLE	Direct Submission	
JOURNAL	Submitted (22-MAY-2000) Departments of Genomic Sciences and Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore Drive, Research Triangle Park, North Carolina 27709-3398, USA	
FEATURES	Location/Qualifiers	
SOURCE	1. 3721 /organism="Staphylococcus epidermidis" /strain="SRI" /db_xref="taxon:1282" /clone="step.1032d02"	
BASE COUNT	1204 a 625 c 495 g 1397 t	
ORIGIN		

Query Match	15.1%	Score 155;	DB 2;	Length 3721;
Best Local Similarity	76.1%;	Pred. No. 4e-19;		
Matches 191;	Conservative	0;	Mismatches 60;	Indels 0;
			Gaps	0.
QY 1	ATGATTTATGCACTTTCACACTTGCACAAATATGAGTACAAAGGTCTTAATCCACC	60		
DB 255	ATGATGATGATGCTATTTCACACTAGCAAAAATGTGTAATGCAACAAAGCTTTAATCCACCA	200		
QY 61	GTTGGCGCTGTTGTAGTTAATGAAGGTAGCATGTTGTGTTGGTGCACACTTTGAGAAAA	120		
DB 199	GTAGATCCGTTGTTGTTTAAANAACGTAGAGATTGTAGGTTTAGGTGACATTTAAAAAAG	140		
QY 121	GGTGCACAGCATGCGGAGGTTCAAGCACTGTGATATGCAACAACAAATGCTGAAGGTGCG	180		
DB 139	GGAGATTAAACATGCCGGAAGTAGCAACCTATTGTGAAAGTGCAGGTTTAAATACCAAGGTGCT	80		
QY 181	ACGATTTTATATAGCTTAGAGCCATGTAGTCAATTTTGTTCAAACACACCCCTGTGTAAAC	240		
DB 79	ACCAATATACGTTTCATTTAGAAACCTTGACACACACATGTTCAACACACACCTTGTTGCAT	20		
QY 241	AAAATTATTGA 251			
DB 19	AAAATCATTTGA 9			

RESULT	6				
AP001512	LOCUS	AP001512	300550 bp	DNA	BCT
	DEFINITION	Bacillus halodurans genomic DNA, section 6/14.			
	ACCESSION	AP001512	BA000004		
	VERSION	AP001512.1	GI:10174030		
	KEYWORDS				
	SOURCE				
	ORGANISM	Bacillus halodurans DNA.			
		Bacillus halodurans			
		Bacteria; Firmicutes; Bacillus/Clostridium group;			
		Bacillus/Staphylococcus group; Bacillus.			
REFERENCE		1 (sites)			
AUTHORS		Takami,H., Nakasone,K., Hirama,C., Takaki,Y., Masui,N., Fujii,F., Nakamura,Y. and Inoue,A.			
		An improved physical and genetic map of the genome of alkaliphilic			
		Bacillus sp.-C-125			
JOURNAL		Extremophiles 3 (1), 21-28 (1999)			
MEDLINE		99184645			
REFERENCE		2 (sites)			
AUTHORS		Takami,H., Nakasone,K., Ogasawara,N., Hirama,C., Nakamura,Y., Masui,N., Fujii,F., Takaki,Y., Inoue,A. and Horikoshi,K.			
		Sequencing of three lambda clones from the genome of alkaliphilic			
		Bacillus sp. strain C-125			
JOURNAL		Extremophiles 3 (1), 29-34 (1999)			
MEDLINE		99184646			
REFERENCE		3 (sites)			
AUTHORS		Takami,H., Takaki,Y., Nakasone,K., Hirama,C., Inoue,A. and Horikoshi,K.			
		Sequence analysis of a 32-kb region including the major ribosomal			
		protein gene clusters from alkaliphilic Bacillus sp. strain C-125			
JOURNAL		Biosci. Biotechnol. Biochem. 63 (2), 455-455 (1999)			
MEDLINE		99209008			
REFERENCE		4 (sites)			
AUTHORS		Takami,H., Masui,N., Nakasone,K. and Horikoshi,K.			
		Replication origin region of the chromosome of alkaliphilic			
		Bacillus halodurans C-125			
JOURNAL		Biosci. Biotechnol. Biochem. 63 (6), 1134-1137 (1999)			
MEDLINE		99356711			
REFERENCE		5 (sites)			
AUTHORS		Takami,H., Takaki,Y., Nakasone,K., Sakiyama,T., Maeno,G., Sasaki,R., Hirama,C., Fujii,F. and Masui,N.			
		Genetic analysis of the chromosome of alkaliphilic Bacillus			
		halodurans C-125			
JOURNAL		Extremophiles 3 (3), 227-233 (1999)			
MEDLINE		99411980			
REFERENCE		6 (sites)			

AUTHORS	Takami, H.
TITLE	Genome analysis of facultatively alkaliphilic Bacillus halodurans C-125
JOURNAL	(10) Extremophiles in deep-sea environments (Ed.):
REFERENCE	: 249-284; Springer-Verlag (1999)
AUTHORS	7 (sites)
TITLE	Takami, H. and Horikoshi, K.
JOURNAL	Reidentification of facultatively alkaliphilic Bacillus sp. C-125 to Bacillus halodurans
REFERENCE	Biosci. Biotechnol. Biochem. 63, 943-945 (1999)
AUTHORS	8 (sites)
TITLE	Takami, H., Nakasone, K., Takaki, Y., Maeno, G., Sasaki, R., Masui, N., Fujii, F., Hirama, C., Nakamura, Y., Ogasawara, N., Kuhara, S. and Horikoshi, K.
JOURNAL	Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and genomic sequence comparison with Bacillus subtilis
REFERENCE	Nucleic Acids Res. 28, 4317-4331 (2000)
AUTHORS	9 (sites)
TITLE	Takami, H. and Horikoshi, K.
JOURNAL	Analysis of the genome of an alkaliphilic Bacillus strain from an industrial point of view
REFERENCE	Extremophiles 4 (2), 99-108 (2000)
AUTHORS	10 (sites)
TITLE	Nakasone, K., Masui, N., Takaki, Y., Sasaki, R., Maeno, G., Sakiyama, T., Hirama, C., Fujii, F. and Takami, H.
JOURNAL	Characterization and comparative study of the rrn operons of alkaliphilic Bacillus halodurans C-125
REFERENCE	Extremophiles 4 (4), 209-214 (2000)
AUTHORS	11 (bases 1 to 300550)
TITLE	Takami, H. and Takaki, Y.
JOURNAL	Direct Submission
REFERENCE	Submitted (22-MAR-2000) Hideyo Takami, Japan Marine Science and Technology Center, Deep-sea Microorganisms Research Group; 2-15 Natsushima, Yokosuka, Kanagawa 237-0061, Japan
AUTHORS	(E-mail: takamih@jamstec.go.jp; URL: <a href="http://www.jamstec.go.jp/Jamstec/e/bio/DEEPSTAR/Research.html">http://www.jamstec.go.jp/Jamstec/e/bio/DEEPSTAR/Research.html</a> ), Tel: 81-468-67-3895, Fax: 81-468-66-6364)
FEATURES	Location/Qualifiers
source	1. 300550
gene	/organism="Bacillus halodurans"
CDS	/db_xref="taxon:86665"
gene	/note="alkaliphile"
CDS	157..816
gene	/gene="BH1414"
CDS	157..816
gene	/gene="BH1414"
CDS	/note="BH1414 unknown"
gene	/codon_start=1
CDS	/transl_table=1
gene	/protein_id="BAB05133.1"
CDS	/db_xref="GI:10174031"
gene	/translation="MPEEGYQQLLEILKLRQGLSFRKIAQALNTVGVKHYRKKH EELDQGGAKPAKGEWQLGSGYVEVETITLPGPTLYAFMPTSEOTRGVSYQDTE KMDQLKRLRLATVYTSILDFGRANRYVDIHLPEKMNMTLIEHVEPNCYTLIDGLVGT PPSAFLPYVASTPIETIPROKNDQYNRKEDYLDMDQGLSEPKMLEHFSITYYEKT K"
gene	829..3600
CDS	/gene="BH1415"
gene	829..3600
CDS	/gene="BH1415"
gene	/note="BH1415 unknown conserved protein in others"
CDS	/codon_start=1
gene	/transl_table=1
CDS	/protein_id="BAB05134.1"
gene	/db_xref="GI:10174032"
CDS	/translation="MREGYGSIVYHAHLPYRHQEDRLERMLPEAMSEYIPLNNA LEKPIKHAAYTISTPPYEMKLSPPVYQTRILNHLNENTBQLKREKRTNQRQNTLV OFYQKREKKAIFLQMDRNLLIGFSRLMEWQCTLMTSAATYHAFPLYLTKREAIRAG

gene  
CDS  
3618..4367  
/gene="mpg"  
/note="BH1416"  
/codon\_start=1  
/transl\_table=11  
/product="mannose-1-phosphate guanylyltransferase"  
/protein\_id="BAB05135.1"  
/db\_xref="GI:10174033"  
/translation="MKGVTLAGKRCRRLKPLTDOIIPKMLPIAGVPCILANGLAHLAAR  
GIRDLVIMVHYLNHMQKAYFODGSKYGRIRIYYVOBDAPLQAGSLKKAERYLDLDFPVV  
MSGDLTITISTOEAIIVFKRONSILMTITKKNAGNONGVYVOTPPNHRVAFRRKPIEM  
DKTRIVLNTGLIYMDPVLISYIPKSGSVLDGKVIPIYLVRKLDITFALBEGGYRWDI  
GSPHYLDANDVNSGVSFVGRPSLSLRD"  
4855..5415  
/gene="BH1417"  
4855..5415  
/gene="BH1417"  
/note="BH1417"  
/codon\_start=1  
/transl\_table=11  
/product="5-formyltetrahydrofolate cyclo-ligase"  
/protein\_id="BAB05136.1"  
/db\_xref="GI:10174034"  
/translation="WEEKKALRNRIKKILRSQPEQRKESQRYADILFLSPENNEAR  
AIGTILISPHVOTELIISRGMEGKQIAYRVKRGKTOMTFEYIHSFVLEDFEFL  
REPKSECRQIDGELDLLPVGAFPDCKGNRLGSGYGDYDFRFLSHYKKTITALYSSQ  
OLVESPEDEDERVQMITVERGYR"  
5412..6200  
/gene="BH1418"  
5412..6200  
/gene="BH1418"  
/note="BH1418"  
unknown conserved protein in others"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAB05137.1"  
/db\_xref="GI:10174035"  
/translation="MIVALLGVLLSVYAPKTKKLTLGSAISAFVVGALISFGLGFR  
GLLLAIIFFPSGFWSRVYQERKENEITAKSTGDKQVNLNAGTGAICALISFIDDD  
PIYLGFEVASTAANADYVASEVGLPARRIILIKKPVAGTSGVSAIGTAAPAA  
GSFIYVVISIPEWMSFSAHHLLFSLTLAFLNGLPGLTAGGVGLYQCPRGCLKET  
ERKIHGNETPKKKGLRFLNDDYVNAICTGTGALFGIYAGILL"  
6226..7248  
/gene="moeb"  
6226..7248  
/gene="moeb"  
/note="moeb"  
/note="BH1419"  
/codon\_start=1  
/transl\_table=11  
/product="molybdopter in biosynthesis"  
/protein\_id="BAB05138.1"  
/db\_xref="GI:10174036"  
/translation="MTDLREYSROMLAPIGKGOOKLONSAYLIVGICALCTVTA  
NHFPVAGHYVMDRODVYVNSLORLILPEPNYRECLPAPVAAOOLKOVNSDIKV  
EGIVADYVENVIHLEMGMDYLDTGTNREFRLINDACRMGLPIYGGVYSGRGT  
AMFLPDKTLCRGTIOBGQSGGTCDTYGVGLPYVDYVASFQALEAKYLTNDQAIR  
HSLMSPVNRHRRHDMNFAAKKANCPTCGTKAPALHPAAHDSVTSLCGRYTVOVHVG  
RPLFSDMEGKGLDRIDVQTPPLVAVQLDEGERLVLEKGRVLYOCTEIVSRKTLTV





Db 159658 ACAGCTTATGTAACATTAGAACGTAATCATTTTGTAAACCCCTCCATGCTGTAAT 159717  
Qy 241 AAAATTATGATTTAGTAAGCAAAAAGTAGTATACCAACAAAACAAATTCGTTAGAC 300  
Db 159718 GCACCTCTTAATCTGAATTAATCGCTGTTCTATTTCTATATATAGATCCATCCITAAA 159777  
Qy 301 ACACATGCTGATGAGA-----CGTTACGGGCGTCAGGCTATTTGAGTTGAATGCGTTGAT 354  
Db 159778 ATTTCTGTAATGGAATACTATATTTTAAAAAACACGATTTTGTGTAAAAACAGGTTTG 159837  
Qy 355 GATGAACGGGCATC---ACATTTATACCAACTTTTAAAGCAAAAACCAACCACTG 411  
Db 159838 CTATCAAAAACATCAAAACATATATATTAAGGGTTTAAAGAAACAAACGATTTT 159897  
Qy 412 CCACAATTACAGTGAAGATGTCAGATTAGATGTTAAACAAGCAATGATATAGA 471  
Db 159898 CCATGATATACAGCTTAAATATAGCATATTCATGATGGAAGAATTCGAATGAGAAACGGC 159957  
Qy 472 CAAGTCATGATGATTTACTTAACAAGAGGTTAAACAAGATGCTATTAAGATTAGACATGCA 531  
Db 159958 GATAGTAAATGATTTACTTCAAAACAGCGCGTCACAGATGTTCAAAATTTGAGCAAAA 160017  
Qy 532 CACGACCACTGTTAATCTGAGACGTCACAGTTGAATTAAGATGATCCACAATATACTACA 591  
Db 160018 AGTTCACTAATTTCACTACACAGTTCTACTATTTTAAAGTGAATVCCCTTATTAACGTA 160077  
Qy 592 CGTATTCAGATGAGAAACCCCTATTAAGTATATTTGCTAA 635  
Db 160078 CGATATTAACAGCTTGATTAACAAACACATATCTATTTTCCCTAA 160121

RESULT 8  
BSRIB 6006 bp DNA BCT 10-FEB-1999  
LOCUS B.subtilis riboflavin biosynthesis operon ribdg, ribb, riba, ribh,  
DEFINITION and ribt genes.  
X51510  
X51510.1 GI:40083  
VERSION riba gene; ribb gene; ribdg gene; ribh gene; riboflavin  
KEYWORDS biosynthesis; riboflavin biosynthesis operon; ribt gene.  
SOURCE Bacillus subtilis.  
ORGANISM Bacillus subtilis.  
Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillaceae;  
Bacillus.  
REFERENCE 1 (bases 1 to 6006)  
AUTHORS Mironov V.N.  
TITLE Direct Submission  
JOURNAL Submitted (23-JAN-1990) Mironov V.N., Centre of Bioengineering,  
Institute of Molecular Biology, Academy of Science, Vavilov Str 32,  
Moscow 117984, U S S R  
REMARK (AA 1-6006)  
REFERENCE 2 (bases 858 to 1246)  
AUTHORS Mironov V.N., Perumov D.A., Kraev A.S., Stepanov A.I. and  
Skirblad K.G.  
TITLE Unusual structure of the regulatory region of the riboflavin  
JOURNAL biosynthesis operon in Bacillus subtilis  
MEDLINE Mol. Biol. (Mosk.) 24 (1), 256-261 (1990)  
REMARK 90271920  
REFERENCE (AA 1-6006)  
AUTHORS 3 (bases 1 to 6006)  
Perumov D.A.  
TITLE Riboflavin operon of Bacillus subtilis: unusual symmetric  
JOURNAL arrangement of the regulatory region  
MEDLINE Mol. Gen. Genet. 233 (3), 483-486 (1992)  
REMARK 92318904  
TITLE Data kindly reviewed (03-SEP-1990) by Mironov V.N.  
COMMENT Location/Qualifiers  
FEATURES  
1. 6006  
/organism="Bacillus subtilis"  
/strain="SHGW"  
/db\_xref="taxon:1423"

CDS  
/map="210 units"  
475..819  
/note="ORF x (AA 1-114)"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAA35877.1"  
/db\_xref="GI:40084"  
/translation="MGRIKRTITLVLLILLAGVYINDIELKDPVPAIGOTLSE  
EEETIOEVKTKIDGSEYHVAENGRKIIIFNCKKLINQDLSIDKEGDKIAYNSKSR  
IDGLIKVAKVND"  
882..910  
/note="promoter region"  
918..5193  
/note="primary transcript"  
1212..2297  
/note="ribdg protein product (AA 1-361)"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAA35878.1"  
/db\_xref="GI:40085"  
/db\_xref="SWISS-PROT:P17618"  
/translation="MEEYMKALDLADAKOEGGTESNPLVGAIVYKDGQVYGCAHLK  
YGRHAEVRAIHAGAHASGADIVTLPECSHGKTPPCAEILIIINSIKRVFAMRP  
NPLVAGRGISMKRERAGILEVREGILADAEKLNKRFLEMTGRLPYVTLKAAASLDKI  
ATSGDSKWTISEARODAOQYRKTHOSILVGVYKADNPSTLCRLPNTKQPVRY  
LDVLSIPEDAKVICDOIAPTWIETFAARAEKKRLSAGVNIFTLETETIOPVLY  
KILAEKINSVYEGGSVAHGSFVEKCFEIIIFYPAKLIIGTHAPSLISGSGFSM  
KDVPLQEPDITDIGNDKILAKPTKE"  
2308..2955  
/note="ribb protein product (AA 1-215)"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAA35879.1"  
/db\_xref="GI:40086"  
/db\_xref="SWISS-PROT:P16440"  
/translation="MFTGIIETGTIESMKKACHAMALITKSKLIEDVHLGDSIAVN  
GICLTVDFTKNOFTVDVMPETVATKATSLNDLTGSKVNLRAAANRFGCHFGVSH  
DGTAEITRIEKNVNAVYDDKMPSTLTKTLVKGSLIVDGSFTLIFGLTETVTSII  
PHRISEITFEKRTIGSKVNECMICKYVYRFLKNAENKNTQGTITKAPLSENGF"  
2970..4166  
/note="riba protein product (AA 1-398)"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAA35880.1"  
/db\_xref="GI:40087"  
/db\_xref="SWISS-PROT:P17620"  
/translation="MEHPTEALDALKGEVYIVVDEDERENEGDFVYALAEHATPEVI  
NFMATRGRLICPTLSEIADRLDHPVHNNDSSHATFTVSIDHRETKTGISAOBR  
SFVQALILDSKSPDFORGHIEPLIAKKGVLKRAHTEAAVDLAECGSGAGYI  
CEIAMEDGTARVPELIEIAKHKQIKMTIKDLIOYKNTLTVREYVDITLPTDGT  
FKYVGTNEVDKGEHVAFFVGDVPCGEPPELVVSLCLDGPNGSGRCGCGQIHA  
LNOIAEGRVLYLHNOEGKIGLTKKRYKLOEGSYDITVEANBALGTPDLPRTNGI  
GAOILNDLGVKNNKLLTNPRKIAGLEGYGLSISERVPIOMEKEHKKYTLQTKNNKL  
GHLHP"  
4199..4663  
/note="ribh protein product (AA 1-154)"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAA35881.1"  
/db\_xref="GI:40088"  
/db\_xref="SWISS-PROT:P11998"  
/translation="MNIIGNLVGTGKIGIVVGRFNDFTSKLSGAEDALLRHGD  
TNDIDVAVGAEIPEPAKMAETKKYDAIITLGYIRGATTHDYVVCYEAAGKGLAQ  
AANTGVPIVFGIVTTENIQALIERAGTKNGKNGVDCAVSAIEMANLNRSFE"  
4776..5150  
/note="ribt protein product (AA 1-124)"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAA35882.1"  
/db\_xref="GI:40089"  
/db\_xref="SWISS-PROT:P17622"  
/translation="MLIRYKSEFKIAMGLSFPMPNEKDLKQLQOTIKDYETDTRDL

CDS

FLKKEEDIVCAIVEKKEKDSVEVETRIHSVNFNSHRHOGIGKOMMALKHLFKYQVLVFN  
ELTOSFEERCCGGODDODISYNN"  
complement(5164, .5688)  
/note="ORF Y (AA 1-174)"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CA35883.1"  
/db\_xref="GI:580916"  
/db\_xref="SWISS-PROT:P17617"  
/translation="MPKAYIDYVEFHATRDYFECHIELEWEKEDPPKKRKYRWG  
FIOLAVAIYHRRONTAGKARLUMANSIILOAEHRAVEDLGDHGRLELQOSYVEOI  
ETVSAYSIILPITDEKLEACRIECCRKKKTWQPSALSNIFLIDKRLRLDRDVIIR  
EREKEIERKRSRD"

BASE COUNT 1891 a 1167 c 1401 g 1547 t  
ORIGIN

Query Match 11.7%; Score 120.4; DB 2; Length 6006;  
Best local similarity 51.0%; Pred. No. 8.3e-13;  
Matches 317; Conservative 0; Mismatches 296; Indels 9; Gaps 1;

OY 1 ATGATTTATGCGATTCACCTTGCAATATGTTACAAAGTCAACAGGTGTTATCCACC 60  
DB 1227 ATGAAGCGCGCTTAGATCTTCGGAAGCAGGCGCAAGAGCAGACCCAAATCCCGCTC 1286  
OY 61 GTTGGCGCTGTTGTAAGTAAGAGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 120  
DB 1287 GTGGCGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1346  
OY 121 GGTGCAAGCATGGGAGGAGTCAACGACTGTTATGCGACACCAAAATGCTGAAGTGG 180  
DB 1347 GGTGAAGCTCAAGAAAGTTGATGCTCAATATGCTGCGACACATGACAGAGGTGGC 1406  
OY 181 AGATTTATATTACGTTAGAGCATGATGTCATTTGTTGTTCAACACCAACCTGTGTTAA 240  
DB 1407 GACATTTAGCTTACACTGCAACCGTGCAGCCATFACGGAAAAACACCGCATGTGAGAA 1466  
OY 241 AAAATTATGATTTGTAAGTACGCAAAAGTATGATACCGCAACAAAGA-----CAAT 291  
DB 1467 TTGATTTATCACTGTGTAAGAAAGTGTGCTGCGGATGAGAGATCTTAATCCGCTT 1526  
OY 292 TCGTTAGACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 351  
DB 1527 GTGGCTGGAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1586  
OY 352 GATGATGAACGGGATCACAATTTATACCAAGACTTTTAAAGCAAAAGCAAAAGCAACTG 411  
DB 1587 CTGGCAGACCAAGCGGAGGCTGAATGAAATTTCTGCACTTATGAGACAGGCGCTT 1646  
OY 412 CCACAATTTACAGTGAAGTATCTGCAAGTTAGATGCTAATCAACAGCGAATGATATGCA 471  
DB 1647 CCGTAGCTACGCTAAAGCGGCTGCCAGCTTGACGGCAAGATAGCTACGACAGGCGT 1706  
OY 472 CAAGCTCAATGATTTACTAACAAGAGTTAAACAAGATGCTATTAAGTTAGACATGCA 531  
DB 1707 GACACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1766  
OY 532 CACGACGACAGTGTTAACGTGAAGACGTACAGTTGATTAAGATGATCCACATATATCTACA 591  
DB 1767 CACCAACGATTTTATGAGTGTGACAGTGAAGCCGACAAATCCGAGCTTAACCTGC 1826  
OY 592 CGTATTTCAAGATGAAAAACC 613  
DB 1827 AGACTGCCGAATGTAACAAAAC 1848

RESULT 9  
LOCUS BACDIA 28206 bp DNA BCT 26-MAY-1995  
DEFINITION Bacillus subtilis spova to sera region.  
ACCESSION L09228  
VERSION L09228.1 GI:410114  
KEYWORDS 3-dehydroquinase dehydratase; aroc gene; diamniopimelate

## SOURCE

decarboxylase; lysa gene; penicillin-binding protein; peptidyl-prolyl isomerase; phosphoglycerate dehydrogenase; pili gene; response regulator; response regulator kinase; riba gene; ribg gene; ribd gene; ribg gene; ribt gene; riboflavin biosynthesis operon; sera gene; signal peptidase; sip5 gene; spoa gene; spovaf gene.  
Bacillus subtilis (strain 168, sub-species Marburg) DNA.

## REFERENCE

1 (bases 1 to 28206)  
Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Bacillus.

## AUTHORS

Yamamoto, J., Shimizu, M. and Yamane, K.  
Molecular cloning and analysis of nucleotide sequence of the Bacillus subtilis lysa gene region using B. subtilis phage vectors and a multi-copy plasmid, pUB110

## TITLE

Agric. Biol. Chem. 55 (6), 1615-1626 (1991)  
91345841

## REFERENCE

2 (bases 1 to 28206)  
Buchanan, C.E. and Ling, M.L.  
Isolation and sequence analysis of dacB, which encodes a sporulation-specific penicillin-binding protein in Bacillus subtilis

## AUTHORS

J. Bacteriol. 174 (6), 1717-1725 (1992)  
92193254

## JOURNAL

3 (bases 1 to 28206)  
Soroikin, A., Zumbstein, E., Azavedo, V., Ehrlich, S.D. and Serror, P.  
The organization of the Bacillus subtilis 168 chromosome region between the spova and sera genetic loci, based on sequence data

## JOURNAL

Mol. Microbiol. 10 (2), 385-395 (1993)  
95020538

## MEDLINE

location/Qualifiers  
1. .28206

## FEATURES

/organism="Bacillus subtilis".  
/strain="168"

## CDS

/db\_xref="taxon:1423"

## gene

1. .1239  
/gene="spovaf"

## CDS

/gene="spovaf"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="AA67472.1"

## CDS

/db\_xref="GI:410115"

## CDS

/translation="VDIVENRLNAOYEKVTLDDETDOVLGLAVIIVEGAFII  
DVSRYPGNPEEPDTEKVRGARDGFENIVNTALRRIRDERLRKVKTKVERSK  
TDLISCTYEDJADDELVEIYEKEIASIDVDLTMADKVEEIVINOSNPPLRYTE  
RVDVAHVHLEGHVILIVDPSVYITPTLFLHVHVAEEVROASVGTPLRMVRFEG  
ILASTREIWEFLFVLQDPLDPMKRGKPKDHPHILIOIPLADGCTPLMAAIIH  
TPPALSTRMGLIAAVLVIGQIAIEVGLSEPEYIIVSLAISTFTTPSYELSLATNRS  
CPHDTKCFISTKRAHRLYSANVYMASIKSLQTPYLMPLIPNGKALMÖVLVTRAKPG  
AKVRPSIVHPNRNROPTNS"

## CDS

1158. .1164  
/gene="spovaf"

## CDS

/citation=11  
1245. .1250

## CDS

/gene="lysa"

## CDS

/note="putative"

## CDS

/gene="lysa"

## CDS

/note="putative"

## CDS

1245. .2664  
/gene="lysa"

## CDS

1266. .1272  
/gene="lysa"

## CDS

/note="putative"

## CDS

1345. .2664  
/gene="lysa"

## CDS

/codon\_start=1  
/transl\_table=11  
/product="diamniopimelate decarboxylase"

## CDS

/protein\_id="AA67473.1"

## CDS

/db\_xref="GI:410117"

## CDS

/translation="MPLHGTSRONQGHLEIGVDALYLAERKGYSTPLYVYDALIRER  
AKSPKAFISAGLKAQYAVASAKAFSSVAMIOLEAEBSISLDVVGSGELLYAVAGCPA  
ERHIFHGNNKREELRALERHICIGYVDNFTALLEDLCKEKGHSIDVLLRTTPGV









TITLE	The complete genome of the hyperthermophilic bacterium Aquifex
JOURNAL	aeolicus
MEDLINE	Nature 392 (6674), 353-358 (1998)
REFERENCE	2 (bases 1 to 14364)
AUTHORS	Decker, G., Warren, P. V., Gaasterland, T., Young, W. G., Lenox, A. L., Graham, D. E., Overbeek, R., Sneed, M. A., Keller, M., Aujay, M., Huber, R., Feldman, R. A., Short, J. M., Olson, G. J., and Swanson, R. V.
TITLE	Submitted
JOURNAL	Direct Submission
COMMENT	Submitted (25-JUN-1997) Diversa Corporation, Genomics, San Diego, CA 92121
FEATURES	Putative indicates no similarity to known proteins Hypothetical indicates similarity to a protein of unknown function.
source	Location/Qualifiers
gene	1..14364
	/organism="Aquifex aeolicus"
	/strain="VF5"
	/db_xref="taxon:63363"
	72..410
	/gene="trpS"
	/note="ag_123"
	72..410
	/gene="trpS"
	/codon_start=1
	/transl_table=11
	/product="ribosomal protein S16"
	/protein_id="AAC06492.1"
	/db_xref="GI:2982872"
	/translation="MAVRIRLAKFGKRHHPIRYIVMDAKSPREGKIYIDLTGYPKR KVLNVPEKYEVLKGVLSHRAKATLMNGLKEVVPBGYEKKRGDYVPEKRE SKSKGGEAA"
	407..646
	/gene="ag_124a"
	407..646
	/gene="ag_124a"
	/codon_start=1
	/transl_table=11
	/product="hypothetical protein"
	/protein_id="AAC06503.1"
	/db_xref="GI:2982883"
	/translation="MSALDIYELTAKELVDNKRKYRTEIEGKTVVLELVDAEL GKVIKGRIRALITLTAIGRKIGKRVLELLE"
	646..1044
	/gene="ag_125"
	646..1044
	/gene="ag_125"
	/codon_start=1
	/transl_table=11
	/product="putative protein"
	/protein_id="AAC06497.1"
	/db_xref="GI:2982877"
	/translation="MKKLTALALSFPLGSGFSFVYAPTLFVKVLRQAGAVVEKV FPVYGLGIIIVGISLFLGRSGKFLSLGIINLLLLQEFIVIPKLHGLKATVVEL FLKYHGMGINLAILLTLGLKVLILFKR"
	1049..1492
	/gene="ag_126"
	1049..1492
	/gene="ag_126"
	/codon_start=1
	/transl_table=11
	/product="putative protein"
	/protein_id="AAC06498.1"
	/db_xref="GI:2982878"
	/translation="MNKEGVKELLISYKNIADLEPINTKLIYKXKLKAQIITLS LDVSKTEEMFKDMLGEGNEVAVEIAKEMPTQNMTERVYMLFFESTSEVLKEPMEG DYIEDKHELISQTLKISLLEKRLDLKEKOKGLKFLIRLIFFR"
	1492..2310
	/gene="ag_127"
	1492..2310
	/gene="ag_127"
	/codon_start=1
	/transl_table=11
	/product="putative protein"
gene	
	/protein_id="AAC06499.1"
	/db_xref="GI:2982879"
	/translation="MLSLFVVLVLAQAFTFGAGIFVIGILSLFYEPRKIVYST VNLGIITSMILFLGKTRPNKRIEPLIAGSAGISAKLIMDEIVKALIGF VFLGIVDFIVOTNSLKFRKASPFVGTGSLGIFGLMGSGPPVYLVNOCRD LNTFKITISPPRATNVVRIIFYLKGTETYSYLLIPAFIAGLGVFAGIYLSRYF TPYFKRTISLSLSGLFLLGKESHDFFKHSVIEESPYGLKSKHPGN"
	complement(2134..3552)
	/gene="ag_128"
	complement(2134..3552)
	/gene="ag_128"
	/codon_start=1
	/transl_table=11
	/product="hypothetical protein"
	/protein_id="AAC06502.1"
	/db_xref="GI:2982882"
	/translation="MORLIIVNPEPLSVIKKVVLSIPIYVNLTVENNISMILV SISPSAANCFSSILMLPITSMLALSYSGNILLAPVAKKRPSPILINGELSF LISLPFLFYGRDVLFLMKVLGASTFVNSLAKELYLPFWELPIGFLNTPFGAYNGA GQKTPMKVAILIMNLTHIGAVYLLNGKFGLPKLQVEAGWGCIASEILAEPIYFEL IFFKRPFLHLLEPKLLEKVRIGTPTPALERAITLSENFVGLFAKGVLAHQ IGLRISISFMIGFGMIASSTLAGQNGARYRGMVAHVTSAHFTALVMSLTLLI LLEPHYLVPEPSROPEVIEMASYVLOIYGISOPAAVASYSGALKGKTHIPLEVN ISSFWLFRIPSPFLKYTHSPILVPMGTFETPAVALFYTVFVKVYKLLQTPQKE DEPEGEDAGDKPLDEVRSKYT"
	3567..4031
	/gene="ribh"
	3567..4031
	/gene="ribh"
	/note="ag_132"
	3567..4031
	/gene="ribh"
	/codon_start=1
	/transl_table=11
	/product="riboflavin synthase beta subunit"
	/protein_id="AAC06489.1"
	/db_xref="GI:2982869"
	/translation="MQIYEGKLTAEGLRPGIVASPFNNALVDRLVEGAIDQIVRGR BEDITLVVPGSWEIIPVAGELARKEDIDAVIATGVLRGATPHDYTASEYSKLAN LSLELRKPIYTFVITADTLEQATIERAGTKHGKWEALSAIEMANLFRSLK"
	4028..4474
	/gene="nusb"
	4028..4474
	/gene="nusb"
	/note="ag_133"
	4028..4474
	/gene="nusb"
	/codon_start=1
	/transl_table=11
	/product="transcription termination nusb"
	/protein_id="AAC06491.1"
	/db_xref="GI:2982871"
	/translation="MYRKGAQDPAFLVLYRRDLGEMPGELFKVVEEKNIKKDAY ETAKLVDPFAVRIEIDSIIEKHLKWSIDRLGVERNALRGVAELIFLKSKEPGR VFIDIVDVKKRYADEKAGKFGVGLSAIYKAYITSSKEKPSLASE"
	4676..5131
	/gene="ag_134"
	4676..5131
	/gene="ag_134"
	/codon_start=1
	/transl_table=11
	/product="putative protein"
	/protein_id="AAC06500.1"
	/db_xref="GI:2982880"
	/translation="MYOVSVIGSSKASSEETFPAYRLCKELAKRNLLVYCGRGVME AVCKGAKEGGITLIGIMPSTYDGHENPVDIKVNGMNNNPVIVASSENVLAAGN YGLTSEIVAYLILGYVVGKTHKVEGEVESLEEMLKVKDFEGGRL"
	5128..6075
	/gene="nuew"
	5128..6075
	/gene="nuew"
	/note="ag_135"
	5128..6075
	/gene="nuew"
	/codon_start=1
	/transl_table=11
	/product="NADH dehydrogenase (ubiquinone)"
	/protein_id="AAC06490.1"
	/db_xref="GI:2982870"

```

/translation="MKVFTGATGCVGRHIVBELNRCIEVHAGVRLSKLERLFCNQ
VKGIYVNFDEKDISIREALGVNDPFIHLGILYEEKKGITREYHAYGHTKMLVEVS
KGNVKKFLEKDEKSIHDEAPSRVHQIKRAAEVINSNTYIFERSIILGEBOKLF
EDMKITIKYIPVVALPDFGNVQFQPDVADVACVAFALNPEIDRIYLCSTKYVT
FKELLDISFYMDRKVLMIPMKPKIMVFAELIYEKIEPEPSSDQIKMMKKNVCGV
LKDEATSNGLKICKREPIREPYEESLKMSLNFIILKGR"
6098..7165
/gene="cpx"
/note="aq_136"
6098..7165
/gene="cpx"
/codon_start=1
/translation=11
/product="cytochrome c peroxidase"
/protein_id="AAC06485.1"
/db_xref="GI:2982865"
/translation="MRKGLIMAVPAGFVAVKKEIDKELLMKAROFKPLPVANEP
ONPTPEKVKLGKMLYYDPRLSGLISCTCHNLARYCDINLPTSLGHWMATGPRA
PTVYNAIHLAOFWIDGRADVEQALQIPVNPLEMATERNVKTLSIDEVLEFK
APFNEKDPVYENIGAKIGAFERTLMTSPSDEFLGNTKALBOEGRGLKTEIEVGC
VACHNGPVGVMNFAFGMITLEYKVTYVYLVGKPAIKVDGFRFGVTKKEEDMEYFK
VPSLRNIEHTYPYEHDSGVMSLEDVAIRIMAKTQLGKELTDQVKDIYAFLLKALTKIP
KHALEVPBLPSDKTPKPEL"
complement(7162..8232)
/gene="ribD1"
/note="aq_138"
complement(7162..8232)
/gene="ribD1"
/codon_start=1
/translation=11
/product="riboflavin specific deaminase"
/protein_id="AAC06487.1"
/db_xref="GI:2982867"
/translation="MIREIDKNYMKLALSLAKRKGTHTPNTVGAIVVYEGKIVGIG
YHEKAKGPHAEVWALGOAGEKAGATLYLDECTHGRPCTDAIRSGIRVYVA
TIDPNPLMSCKGVEKLRNAGIEVDVCEAEALNEDFTYITTEPHTTLMAQTL
DGKATLVLGSSKMTSKESKRVAILRRETVAVLGVNPIKDPHLYTVFVPEKOP
VRILDELEVLPSARVNLTEAPRTIVITKKEKNEKLEKGVOLILKGFINKNL
KKLKELEIMHLEAVGSPRLTSLKEGFDRIYVFIAPKIMGESLIGDGLIRISIES
LKVRKKVENLGGDLVIFKRY"
complement(8229..9173)
/gene="ribP"
Query Match 11.3%; Score 116.4; DB 1; Length 14364;
Best Local Similarity 50.8%; Pred. No. 4.2e-12;
Matches 386; Conservative 0; Mismatches 361; Indels 13; Gaps 4;
QY 1 ATGATATTATCGATTCACCTTGCAATATATGTTACAAGCTCAACAGCGTTATCCACC 60
DB 8205 ATGAAGTTGGACCTCTCTCGCTTAAGAGAGAAAGCGCTACACTCAAAATCCTAACG 8146
QY 61 GTTGCGCTGTTAGTAAATGAAGTAGATTTGTTGTTATGTCACACTTGAGAAA 120
DB 8145 CTGGGACCGGTTCTGCTAAAGAGAAAGAAATCTGCGGATTAGTTATCAGCAAAAAGCA 8086
QY 121 GGTGACAGCATGCGGAGTTCAAGCACTTGATATGCAACAAAAAATGCTGAAGTCCG 180
DB 8085 GGAAGGCTCAAGCTGAAGTATGCTCGTCAGCGGGGAGAAAAGCAAAAGGGGCA 8026
QY 181 ACGATTTATATTAGTTAGAGCCATGTATGTCATTTTGGTTCAACACACCTGTGTTAAC 240
DB 8025 ACCCTTTACGTTACCTTGAGCCGCTGACTACATTCGGAACACTCCCTCGACGAC 7966
QY 241 AAAATTATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 294
DB 7965 CGCATATTAAGAGAGCGGATTAAGAGGTTTGTAGTGCACACCTTGACCAAAATCTTTTG 7906
QY 295 TTAGACACACATGTTAGAGACGCTTACCGGCTCAGCGTATTTGAGGTTGAATGCGTTG-- 352
DB 7905 ATGCTGTGAAAAGGTTGAAAAGGTTGAAAAGGTTGAAAAGGTTGAAAAGGTTGAAAAGG 7846
QY 353 -ATGATTAACGGCATGACATTTATACCAAGACTTTTAAAGCAAAAGCAAAAGCAACTG 411

```

```

DB 7845 TGCAGAGAGAGCGCAGAGAACTCAACGAGATTTTACTTACTACTACGAGAAAGA 7786
QY 412 CCACAAATTAAGTGAAGATATCTGCAAGTTAGATGTTAAACAGGAAATGATATGA 471
DB 7785 CCTTACATTAAGCTTAAGATATGAGCCGACAGCCCTGACGAGAAACCTCGACCGGCA 7726
QY 472 CAAGTCATGATGATTAACCAAGAGTTAAACAGATGCTATTAAGATTAAGACATCA 531
DB 7725 AGCAGTAAGTGAATACCTCAAGGAAAGCGAAGAGTAGCTCATCTTAAAGAGAGAG 7666
QY 532 CACAGCAGTGTACTGAGAGACGTAAGTGAATTAAGTATGATCCAGATTAAGTACCA 591
DB 7665 GCAACGTGCGTTCGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 7606
QY 592 CG--TATTCAGATGAGAAACCTTAAGATGATGATGATGATGATGATGATGATGATGAT 648
DB 7605 AGTTTGTTCACAGGAGAAACCTTAAGATGATGATGATGATGATGATGATGATGATGAT 7546
QY 649 CATTTTAATCAGCAATTTATCAAGATGATGATGATGATGATGATGATGATGATGAT 708
DB 7545 CCACTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7487
QY 709 CCAATTTTAACAGCAATCAACATCAACATCAACATCAACATCAACATCAACATCAAC 748
DB 7486 GGAGAAATGAGAAAGCTGAAAGAACTCAAGAAAGAAAGAGAGATT 7447

RESULT 12
BARIBGENS 4286 bp DNA BCT -07-AUG-1997
LOCUS B. amylioliquefaciens ribB, ribG, ribA, ribH & ribT genes.
DEFINITION X95955
ACCESSION X95955
VERSION X95955.1 GI:1592687
KEYWORDS deaminase; GTP cyclohydrolase II; lumazine synthase; reductase;
ribA gene; ribB gene; ribG gene; ribH gene; riboflavin biosynthesis
operator; Riboflavin synthase; ribT gene.
SOURCE Bacillus amylioliquefaciens.
ORGANISM Bacillus amylioliquefaciens
Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillaceae;
Bacillus.
REFERENCE 1 (bases 1 to 4286)
AUTHORS Gusarov, I.I., Krenova, R.A., Podchernyaev, D.A., Iomantas, Y.V.,
Abalakhina, E.G., Stoinova, N.V., Perumov, D.A. and Kozlov, Y.I.
TITLE Riboflavin biosynthesis genes of Bacillus amylioliquefaciens:
primary structure, arrangement and regulation
JOURNAL Mol. Biol. 31, 446-453 (1997)
REFERENCE 2 (bases 1 to 4286)
AUTHORS Gusarov, I.
TITLE Direct Submission
JOURNAL Submitted (27-FEB-1996) I. Gusarov, GNIgenetika, 1st Dorozhny pr.
1A, 113545, Moscow, RUSSIA
COMMENT On Oct 4, 1996 this sequence version replaced gi:1212772.
FEATURES
source
1..4286
/organism="Bacillus amylioliquefaciens"
/strain="A 50"
/db_xref="taxon:1390"
321..348
promoter
1..4286
/note="riboflavin operon"
653..1768
gene
1..4286
/note="ribD"
653..1768
CDS
1..4286
/gene="ribD"
653..1768
/codon_start=1
/translation=11
/product="deaminase"
/protein_id="CAA65189.1"
/db_xref="GI:1592688"
/db_xref="SWISS-PROT:P70814"
/translation="MEYVNTAIEELARREGOTRPNPLVGAIVVKKRQVYGAHQLQ
YGAHAEVNAIINNAGSLAKCADILYTLDEPCHGKTPPCGLIMKSGIKRVFLAVEDP
NPLVAGKGTMLEAGIEVYTGILLRQDAEELNKMFLHFMRTGLPYVTLLKAAASLDKGT

```

gene  
 CDS  
 1749. .2396  
 /gene="ribb"  
 /codon\_start=1  
 /transl\_table=11  
 /product="riboflavin synthase (a-subunit)"  
 /protein\_id="CA65190.1"  
 /db\_xref="GI:1212774"  
 /db\_xref="SWISS-PROT:O44680"  
 /translation="MFTGIYETCTIOAIKKTGCSMALTTAAKVTSDVRLGDSIAVNW  
 GICLTIVGSDNOFTVDVMPETVATSLNLGSLGSKYVLEKAMSANGFGGHFVSGH  
 DGTLEITRIEKKSNAAVYDLKLSPELTKTLVAKSKIVDGSSTIFGLSDSVVSVAY  
 PHTISETFRTKAVGSIYVIECDMIGKYLRYFLKTKEDTKSNQITTEAFSENGF"  
 2411. .3607  
 /gene="ribb"  
 2411. .3607  
 /gene="ribb"  
 /note="GTP cyclohydrolase II"  
 /codon\_start=1  
 /transl\_table=11  
 /product="3,4-dihydroxy-2-butanone 4-phosphate synthase"  
 /protein\_id="CA65191.1"  
 /db\_xref="GI:1212775"  
 /db\_xref="SWISS-PROT:P51695"  
 /translation="MFPHIEALEALRKGEYIIVVDEEDRENEGDFVALAEHATPEVNV  
 NFMTHGGLICPTLSEDIAGRLDLHPVNDHTSHETAFVSLIDHKLTKTGISAOEVE  
 SETIOALIDESVPGDFDORPGHIFPLLAKKGRGKRAHPEAAVDLAKAGSGADVIY  
 CEINNEGCTMARVPEISETAKSHOLKMTITKDLTEYRNITTLVNRREDVITLPTDFG  
 FRVYGVINEVDGKEHLAFVWGDVPFNSGPVLYVRVSHCLTGDVASHKDCDGPDLAAHA  
 LRQAEGRGVLVLRQEGRCIGLINIKLAVRLQEGVDPEANEAALGFLPDLRNYGIG  
 GAQILRDLGVGVMKLITNNPRKIAGLIRGYGLSISDVPVLQMEASEHNQYLOTKMKKL  
 GHLLHF"  
 3639. .4103  
 /gene="ribb"  
 3639. .4103  
 /gene="ribb"  
 /codon\_start=1  
 /transl\_table=11  
 /product="lumazine synthase (b-subunit)"  
 /protein\_id="CA65192.1"  
 /db\_xref="GI:1212776"  
 /db\_xref="SWISS-PROT:O44681"  
 /translation="MNTIOGNLVGTGIGIVVGRFMEFITSKLSGAEPLIRHGYE  
 SNDIDVAMVPGAEFIEPRAKKMAETKYDVAIVITIGTVIRGATTHDYVCNAAAGTQ  
 AGTATGVYVIFGIYTTETIEQATIERASTKAGNKGADCAVSAIEMANIRSEF"  
 4217. .4286  
 /gene="ribb"  
 4217. .4286  
 /gene="ribb"  
 /codon\_start=1  
 /transl\_table=11  
 /product="reductase"  
 /protein\_id="CA65193.1"  
 /db\_xref="GI:1212777"  
 /db\_xref="SPTREMBL:O44682"  
 /translation="MLIRYKSPFKIMAGLISFMPNE"  
 gene  
 CDS  
 926 c 1102 g 952 t

Query Match	10.9%	Score 112.2:	DB 2:	length 4286;
Best Local Similarity	52.0%;	Pred. No. 2.7e-11;		
Matches 305;	Conservative 0;	Mismatches 273;	Indels 9;	Gaps 2

**QY** 1 ATGGAATTATGGATTCAACTTGCAAATATGGTACAGAAGGTCAAACAGTGTTAATCCACCC 60  
||| ||| | ||||| | ||||| |||  
**Db** 668 ATGAATACGGCAATAGAGCTTCGCCAGACGGGTGAAGACAGACACAGCGCAATCCTCTT 722

OY	61	GTTCCGCTGGTGTGGTAATGAAGGTGGAAATGTTGGTAATTGGTGACCATCTGGAAAA	120
Dd	728	GTCCGCCCGCCTGCTGTGTCAAAAACGACAGATCTGCAGATAGGCGCCACTTCATATC	787
OY	121	GGTGACAAGCATGCGGAGTTCAGCCACTTGATATGGCAACAATAAATGCTGAAGTGGC	180
Dd	788	GGAGAGGCCCATGTGAGAGGTGCACCAGATCATATGGCGAGCGCATCAAAAAGGCGCT	847
OY	181	ACGATTTTATTTACCTTAGCCATGATGTATCTTTTGGTCAACCCACCGTGTATAC	240
Dd	848	GATTTGTATGTGACCTTGAGCCTTGACGCCATTACGGAAAAAACCCGCATGTGCGCAA	907
OY	241	AAAATTTATTTGATTGTAGATAGCAAAAGTAGTATACGCACAAAAG-----ACAATTGC	294
Dd	908	CTGATCATGAATATCGGAGTATAAAAGGGTGTTATTCGCTGTGAAGATCCCAATCCGCTT	967
OY	295	TTAGCACACACATGTGTATGAGACGTTACCGGCTCACCGGTATTGASGTGAAT---CGCTT	351
Dd	968	TGAGCGGGAAAAAGGATACACAATGCTGGAGAAGCGGGGATGAAAGTGAACGGGGCTT	1027
OY	352	GATGATGAGACGGCGCATCACAATTATACCAGACCTTTTTTAAAGCAAAAAGCAAGCACTG	411
Dd	1028	CTCAGCGCAACAGCGAGAGAGCTGAATAAAATGTTTCTTCACCTTCATGCGGACCGSACTT	1087
OY	412	CCACCAATTTACAGTAAAGTATCTGCAAGTTTNGATGTAAACCAAGCGCATATATATGA	471
Dd	1088	CCGTTGTGCACCTTGAAGAAGCGGCCCAAGACCTTTGACGGGAAAAACGCGACTPAAACGGST	1147
OY	472	CAAAAGTCATAGGATTTACTTAACAAAGAGGTTAAACAAGAATGCTATATAGTTAAGACATCA	531
Dd	1148	GACACCAAAATGSATTTACATCAGAGGCGGCAAGGCTTGATGGCCACAGCAATATACGNAATCG	1207
OY	532	CACGACGCACTTTTAACTGGAAGACGCTACAGTGTGAATTAAGATGATCC	578
Dd	1208	CATCAAAAGATTTTGTGTCGAGACGAGCAAGCTCAAGCCAAGCCGATATATCC	1254

RESULT	13
SPNEU1908/c	
LOCUS	
DEFINITION	SPNEU1908 175936 bp DNA HTG 16-DEC-2000
ACCESSION	Streptococcus pneumoniae clone G54, *** SEQUENCING IN PROGRESS ***
VERSION	in ordered pieces.
KEYWORDS	AL449930
SOURCE	AL449930.1 GI:11545155
ORGANISM	HTG: HTGS. PHASE2.
	Streptococcus pneumoniae.
	Streptococcus pneumoniae
	Bacteria: Filicutes; Bacillus/Clostridium group; Streptococcaceae
REFERENCE	Streptococcus.
AUTHORS	1 (bases 1 to 175936)
	Dopazo,J., Mendoza,A., Herrero,J., Caldra,F., Pollist,A.,
	Humbert,Y., Friedl,L., Guerrier,M., Grand-Schenk,E., Gandin,C., de
	Francesco,M., Buell,G., Feger,G., Garcia,E., Peltsch,M. and
	Garcia-Bustos,J.F.
TITLE	Direct Submission
JOURNAL	Submitted (31-OCT-2000) Research Department, Glaxo Wellcome, S.A.,
	Severo Ochoa 2, 28760 Tres Cantos, SPAIN
COMMENT	* NOTE: This is a 'working draft' sequence.
	* This sequence will be replaced
	* by the finished sequence as soon as it is available and
	* the accession number will be preserved.
	Location/Qualifiers
FEATURES	1..175936
source	/organism="Streptococcus pneumoniae"
	/serotype="19F"
	/db_xref="taxon:1313"
	/clone="G54"
BASE COUNT	52994 a 31653 c 38793 g 52447 t 47 others
ORIGIN	

Query Match	10.8%;	Score 11;	DB 84;	Length 175936;
Best Local Similarity	49.0%;	Pred. No. 3.5e-11;		



gene  
/db\_xref="GI:12723932"  
/translation="WQEPMEKVKRRKRVLLITFSVCTVNLISLVLRQGVNLKYS  
VYAOHLPHKVGTVNRYMAVFNNDLAIYIPNANYNVPEGVAVKTSSEGLERFGLK  
INDISVFKGSELSFKPEAGGLDIDFVTPYTNVNSKGEFISVTSSETPKTIK  
TLOYIIIVKERNEMVGFIVSHR"  
4681..5343  
/gene="ykbf"  
/note="L18811"  
4681..5343  
/gene="ykbF"  
/note="EVIDENCE BY GENEMARK BIO16.01 Unknown"  
/codon\_start=1  
/transl\_table=11  
/product="UNKNOWN PROTEIN"  
/protein\_id="AAK05087.1"  
/db\_xref="GI:12723933"  
/translation="WVRRKIVLFVPLVLLILLMTPELSTERNLHAYSIYAOHV  
HRTGDPVCAVIDNLDNYIPELNKISVEPWRVNVPSKGDIDYDFQIYIN  
ELSKRAOLSVGPPOYNPIGPVKYOTIYSTGAGTSFTNFESOCUSFDSRLMT  
ISNIDYKGFYNLRKKTDLISLKNINEMKRDAYSIIIVKRRKVPISIKLQFLYMLNKR  
FS"  
5393..6004  
/gene="yKCA"  
/note="L19523"  
5393..6004  
/gene="yKCA"  
/note="EVIDENCE BY GENEMARK BIO16.01 Unknown"  
/codon\_start=1  
/transl\_table=11  
/product="UNKNOWN PROTEIN"  
/protein\_id="AAK05088.1"  
/db\_xref="GI:12723934"  
/translation="VKKRRGRFLVFLSIIFVIVLLISGYRNLNRYSIYAON  
IPKSGTDPVCAVIFNLSSVYLPHADLSTEIDRKDNQHWQSYNRYTFYNSQDI  
EYHRANODEVSESMYEFENTYVLSIDKNDNIIYRAEFERSGRLSIDIGDSRTIPT  
NKDEILSWKRRKLVGDIVQKRRKPIINQWLYNLNKRFS"  
6001..6678  
/gene="yKCB"  
/note="L20134"  
6001..6678  
/gene="yKCB"  
/note="EVIDENCE BY GENEMARK BIO16.01 Unknown"  
/codon\_start=1  
/transl\_table=11  
/product="UNKNOWN PROTEIN"  
/protein\_id="AAK05089.1"  
/db\_xref="GI:12723935"  
/translation="MSWVKRRGRFRKVPFLPIFVSVLLIIIGYHNNERYSIYA  
ONIPKSGTAPVCAVIDNLDNIIYIPNINERSKRYNMDVNVFSDRQVLEPDEVN  
SLICKTFEDLEGISPLNODKRTISGINNTRKKIYVLFNFTKKLLKEKRVYSNO  
GREKIFTEINTEKGFVSVEYKNNLATISLQOKLYSIIQSOVPLINLOMLYNL  
NYKRRFS"  
6694..7119  
/gene="yKCC"  
/note="L20815"  
6694..7119  
/gene="yKCC"  
/note="EVIDENCE BY GENEMARK BIO16.01 Unknown"  
/codon\_start=1  
/transl\_table=11  
/product="UNKNOWN PROTEIN"  
/protein\_id="AAK05090.1"  
/db\_xref="GI:12723936"  
/translation="MMNRNTOKTDTPPKRRLLPWLIPLVGMIFLSVITLESYSKY  
AKIYHVEKQELDEVTHTINNGKQLVTSQSNFPIYIDIKASECTDIYVTLIVEN  
DYSVYKNDMPKIKWIIEGMSAGIKLITSEHNGLR"  
7651..8739  
/gene="ribG"  
/note="L0163"  
7651..8739  
/gene="ribG"  
/note="EVIDENCE BY HOMOLOG Y BIO02.09 Riboflavin and  
cobalamin. INVOLVED IN RIBOFLAVIN BIOSYNTHESIS. CONVERTS

2,5-DIAMINO-6-(RIBOSYLAMINO)-4(3H)-PYRIMIDINONE  
5'-PHOSPHATE INTO  
5'-AMINO-6-(RIBOSYLAMINO)-2,4(1H,3H)-PYRIMIDINEDIONE  
5'-PHOSPHATE"  
/codon\_start=1  
/transl\_table=11  
/product="riboflavin-specific deaminase (EC 3.5.4.-)"  
/protein\_id="AAK05091.1"  
/db\_xref="GI:12723937"  
/translation="MKDEYVMDLAVLNKGGVNVNPNQVAGALIVKRGRIIGGYH  
EKYGAHAENAFKDCNESPGLATYVLEPCAHQKQPCFEALIKRIRVYIHL  
DPNPLVSGKGIKAMREACIEVSLANVLECEKINKIFYYSGILPYMMKRYTLDG  
KATISAGSKMTGTGRIARQKHODSRNMAIIVETILDDPOLSVLEGNKPIRI  
ICDTHLRSPILASKIVSTAREIPIIATAESPEKQAPFLKAGCEIILVSCLEGRNLK  
ELTKRIGEGIDSIIILEGGTILNALSALAGVQVQVYIAKIFGKREAPSPVEHGII  
EOPNOAVYLEGGQIYTLIEDILLESLL"  
8739..9389  
/gene="ribB"  
/note="L0164"  
8739..9389  
/gene="ribB"  
/note="EVIDENCE BY HOMOLOG Y BIO02.09 Riboflavin and  
cobalamin. FINAL STEPS OF RIBOFLAVIN SYNTHESIS"  
/codon\_start=1  
/transl\_table=11  
/product="ribflavin synthase alpha chain (EC 2.5.1.9)"  
/protein\_id="AAK05092.1"  
/db\_xref="GI:12723938"  
/translation="MTGIIIEIGFVSRVLRQREKFSIRASLIEMDKIGDSTN  
GICLTVSEFNQKEPTADIMPBSLNSNIGTLNVGAKVULERAMPINQFGHIYSGI  
DVGQIILEKKEBNAIWEIETISLMLKRTIEKSGITIDISLIVAQVKSFSYSII  
PHPTKETIMADKKVDIVNLENDILGKRIEKLWDSSESSMKESVTKERLNNFL"  
9398..10594  
/gene="ribA"  
/note="L0165"  
9398..10594  
/gene="ribA"  
/note="EVIDENCE BY HOMOLOG Y BIO02.09 Riboflavin and  
cobalamin. RIBOFLAVIN SYNTHESIS"  
/codon\_start=1  
/transl\_table=11  
/product="GTP cyclohydrolase II / 3,4-dihydroxy-2-butanone

Query Match 10.3%; Score 105.8; DB 1; Length 11167;  
Best Local Similarity 49.5%; Pred. No. 3.7e-10;  
Matches 334; Conservative 0; Mismatches 332; Indels 9; Gaps 2;

QY 1 ATGGATTATGCCAATTCGAATTCGAATATGTACAAAGTCAAAAGCTGTATATCCACCC 60  
DB 7672 ATGGATTATGCAATTCGACTTGCACAAAGAGAGGTGTAAATGTCAACCCCTCA 7731  
QY 61 GTTGGCGTGTGTAGTATATGATAGTAGATTTGTATGTGTCACACTTGAGAAAA 120  
DB 7732 GTTAGAGCCCTTATGTCTCAAGAGGAGCAATATATGTGCAAGTATATGAAAAATAT 7791  
QY 121 GTTGACACAGTCGGAGGCTTCAGCAGCTTGATATGGCACAACAAATCTGAAGTCGC 180  
DB 7792 GGGAAAGCTCAGCGGGAATCAATGCTTTAAGGATTTGAAGCAATCTCCCAAGAGCC 7851  
QY 181 AGCATTTATATAGCTTATAGCAGTATAGTATGTTGTTCAACACCCAGCTGTATAC 240  
DB 7852 ACCGTTATGTGACCTTAAGCCCTGTGCGCAGCAGGGAACACACCCCTGTGTTGAA 7911  
QY 241 AAAATTATGATTTGAATAGCAAAAGTAGTATACGCAACAAAGAC-----AATTGC 294  
DB 7912 GCGATTATAAAAATAGATTTAAGCGTGTATTATCGGCAATTTAGACCCCAATCCCTTG 7971  
QY 295 TTAGACACACATGCTGATGAGAGCTTACGGCTCAGCGTATTTGAGGTGATCGCTGAT 354  
DB 7972 GTTTCAGAAAAGATTTAAGCAATGAGAAGAGCTGAATGAGTTCTCTCAATGTT 8031  
QY 355 GATGAAGCGGCAT---CAATTTATACCAAGACTTTTAAAGCAACCAAGCAAGCACTG 411



```

Db      8032 TTACAAAAAGATGCAAGATTAATAAATTTCTTTTATTTATGTTAGCAAGGCTT 8091
QY      412 CCACAAATACAGTGAAGTATCTGCAAGTTTGTAGTGAACGCAATGATATGGA 471
Db      8092 CCCTATGTGATGATGAAGTATGTAATGCTTTAGACGMAAATTTGGACATCAGCAGGT 8151
QY      472 CAAAGTCAATGATGATTAACAAAGAGGTTTAAACAGATGCTTAAAGTTAAGACATGCA 531
Db      8152 CAATCAAAATGATGATTAAGTGAATTTGCAAGACAGAAAGATCATCAAGACCTTCAAGT 8211
QY      532 CACAGACGAGTGAAGTGAAGAGTACAGTGAATTAATGATGATGATCCACATATACATACA 591
Db      8212 TTTATGCAATTAATGATGATGCTTGAAGACCTTATTTAGATGATCCACATCTCTGTC 8271
QY      592 CGTATTCAGATGGAAGAAACCCCTATAAGTATATTTGCTTAAGTCTGGGATATTTCAAT 651
Db      8272 AGGTGGAAGAAATGGAAGAAATCCGATTAAGATCATTTGTGATCACTTTAAGAAAGTCTT 8331
QY      652 TTTATATCAGCAAAAT 666
Db      8332 TTGGCTTCCAAAGATT 8346

RESULT 15
LOCUS   U32775 18073 bp DNA BCT 29-MAY-1998
DEFINITION Haemophilus influenzae Rd section 90 of 163 of the complete genome.
ACCESSION U32775 L42023
VERSION   U32775.1 GI:1573951
KEYWORDS
SOURCE   Haemophilus influenzae Rd.
          Haemophilus influenzae Rd
          Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
          Haemophilus.
REFERENCE
AUTHORS  1 (bases 1 to 18073)
          Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A.,
          Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J., Dougherty, B.A.,
          Merriek, J.M., McKenney, K., Sutton, G.G., Fitzhugh, W., Fields, C.A.,
          Gocayne, J.D., Scott, J.D., Shiley, R., Liu, L.I., Glodek, A.,
          Kelley, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T., Hedblom, E.,
          Cotton, M.D., Utterback, T., Hanna, M.C., Nguyen, D.T., Saudek, D.M.,
          Brandon, R.C., Fine, L.D., Fritchman, J.L., Fuhmann, J.L.,
          Geoghegan, N.S., Gnehm, C.L., McDonald, L.A., Small, K.V., Fraser, C.M.,
          Smith, R.O. and Venter, J.C.
          Whole-genome random sequencing and assembly of Haemophilus
          influenzae Rd
JOURNAL  Science 269 (5223), 496-512 (1995)
MEDLINE  95350630
REFERENCE
AUTHORS  2 (bases 1 to 18073)
          Tatusov, R.L., Mushegian, A.R., Bork, P., Brown, N.P., Hayes, W.S.,
          Borodovsky, M., Rudd, K.E. and Koonin, E.V.
          Metabolism and evolution of Haemophilus influenzae deduced from a
          whole-genome comparison with Escherichia coli
JOURNAL  Curr Biol. 6 (3), 279-291 (1996)
MEDLINE  96398784
REFERENCE
AUTHORS  3 (bases 1 to 18073)
          White, O., Clayton, R.A., Kerlavage, A.R. and Fleischmann, R.D.
          Direct Submission
JOURNAL  Submitted (25-JUL-1995) The Institute for Genomic Research, 9712
          Medical Center Dr. Rockville, MD 20850, USA
          4 (bases 1 to 18073)
          White, O., Clayton, R.A., Kerlavage, A.R. and Fleischmann, R.D.
          Direct Submission
JOURNAL  Submitted (27-SEP-1997) The Institute for Genomic Research, 9712
          Medical Center Dr. Rockville, MD 20850, USA
          The H. influenzae sequence has been updated by R. Fleischmann. New
          database matches have been assigned, product names have been
          improved, and a number of frame shifts have been corrected. We
          gratefully acknowledge the work of Tatusov et. al. We have
          incorporated their annotation into the /notes fields of the
          corresponding H. influenzae genes
          5 (bases 1 to 18073)
          White, O., Clayton, R.A., Kerlavage, A.R., Fleischmann, R.D.,

```

```

TITLE      Peterson, J., Hickey, E., Dodson, R. and Gwin, M.
JOURNAL    Direct Submission
REMARK     Submitted (28-MAY-1998) The Institute for Genomic Research, 9712
COMMENT    Medical Center Dr. Rockville, MD 20850, USA
FEATURES   The whole genome was shifted by 588 nucleotides for a new start
SOURCE     On Sep 30, 1996 this sequence version replaced gi:1221677.
           Location/Qualifiers
             1..18073
               /organism="Haemophilus influenzae Rd"
               /db_xref="taxon:71421"
               complement(146..1327)
               /gene="HI0929"
               complement(146..1327)
               /note="similar to GB:M77129 SP:P24196 PID:146678
               PID:882568 GB:U00096 percent identity: 41.10; identified
               by sequence similarity; putative"
               /codon_start=1
               /product="conserved hypothetical protein"
               /protein_id="AAC22589.1"
               /db_xref="GI:1573952"
               /translation="MKRVTFPTPRPMVOQLLVGFDYNYLPSSDSHRSQDVATPE
               TLAIDRIEDTTEBLHSMCLDFADEIKAGDYENHETELQKOLIEFSWRNDEPYLYG
               RFDGVDGNLMKMEYNADPTSLLEAAVQWOMLQIEGLKHRDQNMTHBELIKHF
               QFLKQSGKTDPHLSAMODAGRDGCVNVDLADVAVNAGNHIQLAVEDIGYSETRK
               FVDLNDQPIEMLEFKLYPLEWLSHAERARHTTAEETRIEPAKMLSNKALAKMAR
               YPNHPLHLPAYETFPPELPKDLMSWVKRPILGREGAVNVEYERKNNGVFAKSGSEHTE
               YKNSGYVYQAKFELPFDGMYPIIGSNVGDVACGMLREDFTAVTGNDSHFTPHYFV
               P"
               complement(1328..1948)
               /gene="HI0930"
               complement(1328..1948)
               /gene="HI0930"
               /note="hypothetical protein; identified by Genemark;
               putative"
               /codon_start=1
               /product="H. influenzae predicted coding region HI0930"
               /protein_id="AAC22600.1"
               /db_xref="GI:1573966"
               /translation="MAKKNHSTIKIENNEKNOILVSLTALLGGCSEBOQVDYQ
               SLDDCLADMKIKELCADKNTESTQATTAOQGLINIRDNQNASAYKNPAENNAQ
               ANQSENRAESTPTKAESTDPSLIGAGVGYMAARLAISFLPSYHPGNRAVTTPTG
               QVQVQPTNRVSGPMLVKGNAGSMNSKPVSRGCFSSPNKTHRSSG"
               complement(1921..2406)
               /gene="HI0931"
               complement(1921..2406)
               /gene="HI0931"
               /note="hypothetical protein; identified by Genemark;
               putative"
               /codon_start=1
               /product="H. influenzae predicted coding region HI0931"
               /protein_id="AAC22601.1"
               /db_xref="GI:1573967"
               /translation="MKSHTLYSAVELYIFGKRMHLHLEDEFLGPKROGTHS
               TVEHFVAWEEDDGGENGTFYALDMRONADPIYMDLHVYVDSSTNHHARKLEICDE
               SGVALLLLNGYHAYVDFARLVGYNSKTIPODLSMNTREITTKQAEQMLGVKTI
               R"
               complement(2424..3734)
               /gene="HI0932"
               complement(2424..3734)
               /gene="HI0932"
               /note="similar to GB:M12843 SP:P08324 GB:X82400 PID:147479
               PID:563868 percent identity: 85.02; identified by sequence
               similarity; putative"
               /codon_start=1
               /product="enolase (eno)"
               /protein_id="AAC22590.1"
               /db_xref="GI:1573953"
               /translation="MAKIVKVGIREIISDKNPTVEAEVHLXGPFVGLAABSGASTG

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2001, 11:30:39 ; Search time 109.57 Seconds  
(without alignments)  
5482.440 Million cell updates/sec

Title: US-09-376-633-1

Perfect score: 1029  
Sequence: 1 ATGATTAATGCGATTCAACT.....TAACCTTACGAAGAAGTGA 1029

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 678276 segs, 291890651 residues

Total number of hits satisfying chosen parameters: 1356552

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_0401:\*

- 1: /SIDs2/gcgdata/geneseq/geneseq/NA1980.DAT:\*
- 2: /SIDs2/gcgdata/geneseq/geneseq/NA1981.DAT:\*
- 3: /SIDs2/gcgdata/geneseq/geneseq/NA1982.DAT:\*
- 4: /SIDs2/gcgdata/geneseq/geneseq/NA1983.DAT:\*
- 5: /SIDs2/gcgdata/geneseq/geneseq/NA1984.DAT:\*
- 6: /SIDs2/gcgdata/geneseq/geneseq/NA1985.DAT:\*
- 7: /SIDs2/gcgdata/geneseq/geneseq/NA1986.DAT:\*
- 8: /SIDs2/gcgdata/geneseq/geneseq/NA1987.DAT:\*
- 9: /SIDs2/gcgdata/geneseq/geneseq/NA1988.DAT:\*
- 10: /SIDs2/gcgdata/geneseq/geneseq/NA1989.DAT:\*
- 11: /SIDs2/gcgdata/geneseq/geneseq/NA1990.DAT:\*
- 12: /SIDs2/gcgdata/geneseq/geneseq/NA1991.DAT:\*
- 13: /SIDs2/gcgdata/geneseq/geneseq/NA1992.DAT:\*
- 14: /SIDs2/gcgdata/geneseq/geneseq/NA1993.DAT:\*
- 15: /SIDs2/gcgdata/geneseq/geneseq/NA1994.DAT:\*
- 16: /SIDs2/gcgdata/geneseq/geneseq/NA1995.DAT:\*
- 17: /SIDs2/gcgdata/geneseq/geneseq/NA1996.DAT:\*
- 18: /SIDs2/gcgdata/geneseq/geneseq/NA1997.DAT:\*
- 19: /SIDs2/gcgdata/geneseq/geneseq/NA1998.DAT:\*
- 20: /SIDs2/gcgdata/geneseq/geneseq/NA1999.DAT:\*
- 21: /SIDs2/gcgdata/geneseq/geneseq/NA2000.DAT:\*
- 22: /SIDs2/gcgdata/geneseq/geneseq/NA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1029	100.0	1029	20	X59909
2	1029	100.0	3336	20	X59913
3	1029	100.0	3336	20	X77007
4	1029	100.0	3336	20	X60277
5	1029	100.0	3336	20	X60830
6	1000	97.2	7588	18	V74533
7	352.2	34.2	1269	18	T84023
8	352.2	34.2	1269	18	X59910
9	120.4	11.7	5567	12	O10165
10	120.4	11.7	5567	20	X81947
11	120.4	11.7	5567	20	X81948

12	120.4	11.7	5567	20	V71600
13	120.4	11.7	5567	21	Z93752
14	114.2	11.1	14273	19	V52173
15	112.6	10.9	1101	20	X77179
16	112.6	10.9	3417	20	X77183
17	112.6	10.9	3417	20	X77574
18	112.6	10.9	3417	20	X79558
19	112.6	10.9	3417	20	V72454
20	77.6	7.5	546	19	Z96259
21	76.2	7.4	1230	20	X83297
22	76.2	7.4	4312	19	V27196
23	76.2	7.4	4312	20	X82299
24	75.2	7.3	1512	21	C47216
25	75.2	7.3	1514	21	C47812
26	65.4	6.4	56485	21	A81476
27	65.4	6.4	34980	21	F21612
28	65.4	6.4	837096	21	A81489
29	54.4	5.3	1686	16	O87587
30	51.2	5.0	3095	11	O03875
31	47.8	4.6	9408	21	A70157
32	47.8	4.6	9636	15	O67190
33	47.2	4.6	1826	19	V29477
34	47.2	4.6	1826	19	V37413
35	46.6	4.5	6124	11	O03568
36	46.4	4.5	5589	15	O68195
37	45	4.4	3582	21	A70241
38	44	4.3	1864	8	N71405
39	43.8	4.3	1974	21	A70261
40	43.4	4.2	1998	21	A70212
41	43.4	4.2	19124	18	T72882
42	43.4	4.2	19124	21	Z96287
43	43.2	4.2	933	20	X61528
44	43.2	4.2	1008	20	X61527
45	43.2	4.2	910715	20	X20248

## ALIGNMENTS

RESULT	ID	Sequence	Location/Qualifiers
1	X59909	standard; DNA; 1029 BP.	
XX	AC	X59909;	
XX	DT	04-AUG-1999 (first entry)	
XX	DE	Pyrimidine deaminase and pyrimidine reductase (ribg) DNA.	
XX	XX		
XX	KM	Pyrimidine deaminase; pyrimidine reductase; ribg; antagonist;	
XX	KM	agonist; antimicrobial; antibiotic; Helicobacter pylori infection;	
XX	KW	H. pylori-induced cancer; gastrointestinal carcinoma; gastric ulcer;	
XX	KW	gastritis; ss.	
XX	XX		
XX	OS	Staphylococcus aureus.	
XX	XX		
XX	XX		
XX	PD	03-JUN-1999.	
XX	XX		
XX	PF	23-NOV-1998; 98MO-US24857.	
XX	XX		
XX	PR	25-NOV-1997; 97US-0978456.	
XX	XX		
XX	PA	(SMIK ) SMITHKLINE BEECHAM CORP.	
XX	XX		
XX	PI	Black MT, Burnham MKR, Fedon JC, Hodgson JE, Knowles DIC;	
PI	PI	Lonetto MA, Nicholas RO, Palmer LM, Pratt JM, Reichard RW;	

PI Rosenberg M, Trani CM, Ward JM, Warren RJ.  
 XX WPI: 1999-347572/29.  
 DR P-PSDB: Y15913.  
 XX  
 PT New Staphylococcus aureus polypeptide and polynucleotide useful in  
 the treatment of gastric ulcer and gastritis  
 XX  
 PS Claim 6: Page 5; 48pp; English.  
 CC The present sequence encodes a Staphylococcus aureus pyrimidine  
 CC deaminase and pyrimidine reductase (ribg) polypeptide. Staphylococcus  
 CC aureus ribg and its antagonists are used to treat individuals in  
 CC need of them. Disease related to expression or activity of ribg can  
 CC be determined by analysing the nucleic acid sequence encoding ribg  
 CC or detecting the ribg polypeptide in a sample. ribg can also be used  
 CC to identify antagonists or agonists. ribg, or its related nucleic acid,  
 CC also has use as a vaccine to induce an immunological response in an  
 CC animal. Antimicrobial compounds (e.g. agonists and antagonists of ribg),  
 CC especially broad-spectrum antibiotics, may be of use in the treatment  
 CC of Helicobacter pylori infection. This should decrease the advent of  
 CC H. pylori-induced cancers, such as gastrointestinal carcinoma. The  
 CC treatment should also cure gastric ulcers and gastritis.  
 XX  
 SQ Sequence 1029 BP; 376 A; 162 C; 198 G; 293 T; 0 other;

Query Match 100.0%; Score 1029; DB 20; Length 1029;  
 Best Local Similarity 100.0%; Pred. No. 5.6e-230;  
 Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGGATTATGCGATTCACCTTCGAATATGTACAAAGTCGAAGGTTATATCCACC 60  
 DB 1 atggattatgcgatccaacttgcgaatatacgtacaaagtcgaagtgatccacc 60  
 OY 61 GTTGCGCGTGTGTAGTATGAGTAGATTGTTGATTGTCACACTTGAGAAAA 120  
 DB 61 gttgcgcgctgtgtatgataagtagattgttgattgtgcacacttgagaaaa 120  
 OY 121 GGTGCAAGCATGCGAGGTTACAGCATTGATGCGCAACAAATGCTGAAGGTGG 180  
 DB 121 ggtgcaagcatgcgaggttacagcattgatgcgcaacaaatgctgaaggtgg 180  
 OY 122 ggtgcaagcatgcgaggttacagcattgatgcgcaacaaatgctgaaggtgg 180  
 DB 122 ggtgcaagcatgcgaggttacagcattgatgcgcaacaaatgctgaaggtgg 180  
 OY 181 ACGATTATATTAAGTACGATGAGGATGATGTTGTTGTTCAACACCCGTGTATAC 240  
 DB 181 acgattatattacgtagagcagtagtgcatttgcgtcaacaccgccgtgtaac 240  
 OY 241 AAAATTATGATTGATGATGATGATGATGATGATGATGATGATGATGATGATG 300  
 DB 241 aaaattatgattgattgattgattgattgattgattgattgattgattgattg 300  
 OY 301 ACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
 DB 301 acacatgattgattgattgattgattgattgattgattgattgattgattgattg 360  
 OY 361 CGGCGATCACAATTTTACCAAGACTTTTAAAGCAAAAGCAAGCAACTGCCACAAT 420  
 DB 361 cggcgatcacaaatttaccagaactttttaaagcaaaagcaaacgaactgcacaat 420  
 OY 421 ACAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480  
 DB 421 acagtgaagatgattgattgattgattgattgattgattgattgattgattgattg 480  
 OY 481 TGGATTACTAACAAAGAGGTTAAACAGATGCTATTAAGTTAAGATGACGACGCA 540  
 DB 481 tggattactaacaaagaggttaaacagatgcttataagtttaagatgacgacgca 540  
 OY 541 GTGTTAACTGAGAGCGATGATGATGATGATGATGATGATGATGATGATGATG 600  
 DB 541 gtgtaaactgagagcgatgattgattgattgattgattgattgattgattgattgattg 600  
 OY 601 GATGAAAAAACCTATTAAGTATGTTAGTGGGAAATATTTATATTCG 660  
 DB 601 gatgaaaaaacctatttaagtattgagggaatattttatatttcg 660

DB 601 gatgaaaaaacccataaagtaatatgtcgaagtcgtggaatatattcattatcag 660  
 OY 661 CAAATTTATCAAGATGATCAACACCAATTTGATATATATCAAAATTCATTAACA 720  
 DB 661 caaatttatcaagatgatacacaacatttgatatacaccgaataatccaatttaaca 720  
 OY 721 AGCAATCAAAACACATTTTAAATTTATTTACTTGAAGCTTGTGATTTAACAACAT 780  
 DB 721 agcaatcaaaaacacattttaaatttacttggaagcttgattgatttaacaacattc 780  
 OY 781 CCAATTTATTAATAAGAGAGATGGAATTTGCTAGTCGAGGACGTCACACCTACT 840  
 DB 781 ccaatttatataaagaagagtggaacttgcagtcgagcgagtcacacctact 840  
 OY 841 TCAGATTTCCATCATATATATATGATGAATTTATTTATGATGATGATGATGAT 900  
 DB 841 tcgaatttcctcatctattatagatgattatctctattatagcccggaataat 900  
 OY 901 GCGGATCTGCAATTTATCAATTTTATCAACAAATGATGATGATGATGATGATG 960  
 DB 901 gcggatctgcgaatttatcaattttatacacaatgatgattgattgattgattgattg 960  
 OY 961 AACCAATTTGAATGTTGATGATGATGATGATGATGATGATGATGATGATGATG 1020  
 DB 961 aaccaatttgaatgtgtatccgagttatgataatcaacaatgatgattgattgattgattg 1020  
 OY 1021 AAGAGTGA 1029  
 DB 1021 aagaagtga 1029

RESULT 2  
 ID X59913 standard; DNA; 3336 BP.  
 XX  
 AC X59913;  
 XX  
 DT 04-AUG-1999 (first entry)  
 XX  
 DE Riboflavin biosynthesis operon containing the ribg gene.  
 XX  
 KW Pyrimidine deaminase; pyrimidine reductase; ribg; antagonist;  
 KW agonist; antimicrobial; antibiotic; Helicobacter pylori infection;  
 KW H. pylori-induced cancer; gastrointestinal carcinoma; gastric ulcer;  
 KW gastritis; ss.  
 XX  
 OS Staphylococcus aureus.  
 XX  
 PN W09926475-A1.  
 XX  
 PD 03-JUN-1999.  
 XX  
 PF 23-NOV-1998; 98WO-US24857.  
 XX  
 PR 25-NOV-1997; 97US-0978456.  
 XX  
 XX (SMIK ) SMITHKLINE BEECHAM CORP.  
 XX  
 PI Black MT, Burnham MKR, Fedon JC, Hodgson JE, Knowles DJC;  
 PI Loeblein MA, Nicholas RO, Palmer LM, Pratt JM, Reichard RW;  
 PI Rosenberg M, Trani CM, Ward JM, Warren RJ.  
 DR WPI: 1999-347572/29.  
 XX  
 PT New Staphylococcus aureus polypeptide and polynucleotide useful in  
 the treatment of gastric ulcer and gastritis  
 XX  
 PS Example 2: Page 30-31; 48pp; English.  
 CC The present sequence represents the staphylococcus aureus  
 CC riboflavin biosynthesis operon containing the pyrimidine  
 CC deaminase and pyrimidine reductase (ribg) gene. Staphylococcus  
 CC aureus ribg and its antagonists are used to treat individuals in

CC need of them. Disease related to expression or activity of ribg can  
 CC be determined by analysing the nucleic acid sequence encoding ribg  
 CC or detecting the ribg polypeptide in a sample. ribg can also be used  
 CC to identify antagonists or agonists. ribg, or its related nucleic acid,  
 CC also has use as a vaccine to induce an immunological response in an  
 CC animal. Antimicrobial compounds (e.g. agonists and antagonists of ribg),  
 CC especially broad-spectrum antibiotics, may be of use in the treatment  
 CC of Helicobacter pylori infection. This should decrease the advent of  
 CC H. pylori-induced cancers, such as gastrointestinal carcinoma. The  
 CC treatment should also cure gastric ulcers and gastritis.

XX  
 XX  
 SQ Sequence 3336 BP; 1202 A; 460 C; 677 G; 997 T; 0 other;

Query Match 100.0%; Score 1029; DB 20; Length 3336;  
 Best Local Similarity 100.0%; Pred. No. 7.6e-230;  
 Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGATTATTCGCAATTCACCTTGCATAATATGCTACAGGTCAACAGGTGTTATCCACC 60  
 DB 1 atgattatcgcatactcaactcgaatatagtacaagtgacaacaggtgttaaccacc 60  
 QY 61 GTTGGCGCTGTGTAGTAAAGAGGAGGATTGTTGGTGTGACACACTGTGAGAAA 120  
 DB 61 gtggtcgctgtgtagtaagaggaagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 120  
 QY 121 GGTGACAGGATCGGAGGTTCAAGCACTTGATATGCGACACAAATGCTGAAGTGC 180  
 DB 121 ggtgacaggaatcgagggtgtcaagcacttgatatagtcacacaatgtctgaagtgcg 180  
 QY 181 AGCATTTATTTACGTTAGGCCATGTAGTATTGGTTCAACACCACCTGTGTTAAC 240  
 DB 181 acgattatattacgtttagagccatgtagtcatttggttcaacacacacccctgttaac 240  
 QY 241 AAAATTATTGATTGTAAGATAGCAAAAAGTATACGCAAAAAGCAATTTGTTAGAC 300  
 DB 241 aaattattgattgttaagatagcaaaaagtatagcacaacaaagcaaatctgttagac 300  
 QY 301 ACACATGTGTATGAGAGCTTACGGGCTCACGGTATTGAGGTTGAATGCGTTGATGAA 360  
 DB 301 acacatgtgtatgagagcttaccggctcacggatttgaggttgatgaggtgtgtgtgtgt 360  
 QY 361 CGGGCATCACAATTATACCAAGCTTTTAAAGCAAAAGCAAGCAACCTGCCAATTT 420  
 DB 361 cgggcattcacaattataccagctttttaaagcaaaaagcaaacctgcccacaattt 420  
 QY 421 ACAGTAAAGTATCTCAAGATTAGATGTAACCAAGCGATGATATATGACAAAGTCAA 480  
 DB 421 acagttaaagtatctcaagatttagatgtaacaaagcgatgataatgacaaagtc 480  
 QY 481 TGGATTACTAACAAAGAGTTAAACAAGATGCTATTAAGTAAACATCGACAGCAGCA 540  
 DB 481 tggattacttaacaaagagtttaacaagatgcttataagtttaagtcagtcagcagca 540  
 QY 541 GTTTTAACTGGAAGAGTACAGTTGAATTGATGATCCAAATATCTACAGTATTTCAA 600  
 DB 541 gtttttaactggaagagtagctgaattgataagtcacaaatattactaactaactca 600  
 QY 601 GATGAAAAAACCCTTAATAAGTAATTTGTAAGCTGGAATTTATTTTAAATCAG 660  
 DB 601 gatgaaaaaacccttaataagtaatttgaagctggagatattcattttaaactcag 660  
 QY 661 CAAATTTATCAGATGATCAACACCAATTTGATATATCTGAATAATCCAAATTTAACA 720  
 DB 661 caaatattatcagatgatcaacacaaatttgatatatactgaaatccaaatttaaca 720  
 QY 721 AGCAATCAAAACACATATTTGAATTTTACTTGAAGCTTGTGATTTAAACAATTTCTT 780  
 DB 721 agcaatcaaaacacatatttgaatttacttacttgaagcttgtgtatataaacaattctt 780  
 QY 781 CACAATTTATATAAAGAGAGTGTGGAATTTGCTAGTCAGAGCAGCTCCAACACACTACT 840  
 DB 781 cacaatttatataaagagaggttggaacttctgtagtcgaggaagcaggtccaaactact 840

QY 841 TCAGATTCGCATCTATATATAGATGAATTTATCTTATATGCCCGGAATTAATT 900  
 DB 841 tcagatctcccatctatataatagatgaatttacttctataatgcccggaattaatc 900  
 QY 901 GGGGATCTGGAATTTATCAATTTTATCAACAACAATATGATGATGATGCCAGATCG 960  
 DB 901 ggggatctggaatttatcaattttatcaacaacaatgatgatgataccagatcg 960  
 QY 961 AACCAATTTGAATTTGTCATTCGAGTTTAAATCAAAATGTAATTAACCTTACGA 1020  
 DB 961 aaccaattgaaattgttcatccagagttatlaaatacaaatgtlaaattacattaca 1020  
 QY 1021 AAGAAGTGA 1029  
 DB 1021 aagaagtgga 1029

RESULT 3  
 X77007  
 ID X77007 standard; DNA; 3336 BP.  
 XX  
 AC X77007;  
 XX  
 DT 06-AUG-1999 (first entry)  
 XX  
 DE S. aureus riboflavin biosynthesis operon.  
 XX  
 KW GMP cyclohydrolase II; riba; diagnosis; vaccine; bacterial adhesion;  
 KW Helicobacter pylori infection; gastrointestinal carcinoma; gastric ulcer;  
 KW gastritis; wound healing; infection; drug screening; therapy;  
 KW riboflavin biosynthesis operon; ss.  
 XX  
 OS Staphylococcus aureus.  
 XX  
 PN W09926646-A1.  
 XX  
 PD 03-JUN-1999.  
 XX  
 PF 24-NOV-1998; 98MO-US25068.  
 XX  
 PR 25-NOV-1997; 97US-0977554.  
 XX  
 PA (SMIK ) SMITHKLINE BEECHAM CORP.  
 PI Black MT, Burnham MKR, Depouck C, Fedon JC, Hodgson JE;  
 PI Jaworski DD, Knowles DJC, Lonetto MA, Mooney J, Nicholas RO;  
 PI Palmer LM, Pratt JM, Reichard RW, Rosenberg M, Traini CM;  
 PI Wang M, Ward JM, Warren RL, Zhong Y;  
 XX  
 DR WPI; 1999-357736/30.  
 XX  
 PT New Staphylococcus aureus riba polypeptide - and methods for using  
 PT the polypeptide to screen for antibacterials  
 XX  
 PS Claim 24: Page 36-37; 53pp; English.  
 XX

This sequence represents the S. aureus riboflavin biosynthesis operon, which contains the riba coding sequence of the invention. riba belongs to the GTP cyclohydrolase II family. The polynucleotides and polypeptides may be employed as research reagents and material for the discovery of treatments and diagnostics for diseases, particularly human diseases. They can be used for diagnosis of the disease and staging of disease, and as reagents in differential screening methods. The polynucleotides may be used as a source for hybridisation probes, and for screening of genetic mutations, serotype, and identification, and for organism chromosome identification. The polypeptides can be used to produce antibodies. The polypeptides can also be used in vaccine formulations, and to identify agonists and antagonists. These are used to prevent, inhibit or treat diseases, particularly of Helicobacter pylori infections, such as gastrointestinal carcinoma. Gastric ulcers and gastritis may also be treated. The polypeptides can also be used to treat wounds and in-dwelling devices to prevent bacterial adhesion and infection, and to



Sequence 3336 BP; 1202 A; 459 C; 678 G; 997 T; 0 other;

Query Match	100.0%	Score 1029;	DB 20;	Length 3336;
Best Local Similarity	100.0%	Pred. No. 7.6e-230;		
Matches 1029; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0

QY	1	ATGCAATTATGCGATTCACAACTTGCAGAAATATGCTTCAAGTCCCAACAGAGTTTAAATCCACC	60
Db	1	atgcatattatgcatcttcaactctgcgaataatgatacgaagctccaacacgggttataccacc	60
QY	61	GTTCGCCCTGTGTGATTAAATGAGGTAGGATTGTTCATTGTGCACACTTGGAGAAA	120
Db	61	gttcgcccgtgtgtgattaaatgaggttaggattgttcattgtgcacaccttggagaaa	120
QY	121	GGTCACACAGATGGGAGGTTTCAAGCACTTGATATGCGCAACCAAAATTCCTCAAGTCCG	180
Db	121	ggtcacacagatgggaggtttcaagcaacttgatacgcaacaacaaatgctgaaggctcg	180
QY	181	ACGATTTATATTACGTTTAGAGCCCATGTAGTCATTTTGGTTCAACACACCACCTGTCTTAC	240
Db	181	acgatttatactacgttagagccatgtagtcatcttggttcaacaccacctgtctaac	240
QY	241	AAATATTATTGATTCTTAAGATAGCAAAAAGTATGTAACGCAACAAAAGACATTTCTTAGAC	300
Db	241	aaatatctatcttaagatagcaaaaaagtatactacgcaacaaaagacattctgttagac	300
QY	301	ACACATGCTATGAGACGTTACCGGCTCCACGGATTGAGTTGAATGCGTTTATCATATGAA	360
Db	301	acacatgtgtatgagacgttaccggtctcaagctatctgaagttcgttagatgtaga	360
QY	361	CGGGCATACAAATTATTAACCAAGACTTTTAAAGCAAAAAGCAACCACTGCCACAATT	420
Db	361	cgggcatacaaatattataccaagactttttaaagcaaaagcaagcaactgccaactt	420
QY	421	ACACTGAAATATCTGCAACTTTAGATNGTAAACGAACGCAATGATATATGACAAAGCTCAA	480
Db	421	acagtgaagaatcttcgtaagcttagatggctaaacgaagcaatgataatggaagaagctcaa	480
QY	481	TGGATTACTTAACAAGAGGTTAAACAAGATGTCTATAAATTAAAGATGACACACACGA	540
Db	481	tggattacttaacaagaagggttaacaagaatgtctataagtttaagatcatcgacaacgca	540
QY	541	GTGTTAACTGGAAGACGCTACAGTTGAAATTAGATGATCCAAATATATACACGTATTTCAA	600
Db	541	gtgtaactcvggaagacgttaacgttgaattagatgatacgaataatactaacagtaactcaa	600
QY	601	GATGCAAAAAAACCCTATTAATAAGTAATATTTGTCATACTCGGCAATATTCATTTTAATAG	660
Db	601	gatgcaaaaaaacctataaagaataatgctctaagcttcgggaataatcatcttaataag	660
QY	661	CAAAATTATCAAGATGAATCAACACCAATTTGGATATATCTGAAAATCCAAATTTAACA	720
Db	661	caaatattacaagatgatacaacaacatlttgatatactacgaaatccaaatttaaca	720
QY	721	AGCAATCAAAACACATATTGAAATTTATTTTACTTGGAATCTGTGATTTTAAACAATTTCTT	780
Db	721	agcaatcaaaacacataltgaaatatttacttctgaaagctcttgatalttaacaacaatcttc	780
QY	781	CACAAATTATTAATAAAGAGAGAGTTGGAACCTTTGCTAGTGCAGACGACGCTCAACCACTACT	840
Db	781	cacaattataataaagaagaggtctgaactcttgctagtcagcgagctgcacaacactact	840
QY	841	TCGAAATTTCACATCTATTATATAGATGAATTTATTCTGTATTATGCCCGCAATTAATT	900
Db	841	tcaaatattcccatctattataatagatgaattatctctataatgccccggaattaat	900
QY	901	GGCGATCTCGGAATTTTCATTTTATTCAAACAAGATGATTTAGATGAGATACCAATCGG	960
Db	901	ggcgatctcggaaattatactcaatttatccaacaatgatagtagatcgagataccagatcg	960
QY	961	AACCAATTTGAAATGTCTTCATTCGAGTTATTTAAATCAAAATGTTAAATTAACATTTAGCA	1020
Db	961	aaccaattttgaaatgtcttcattccgagttattttaaatcaaaatgtttaaattaacatttagca	1020

Db 961 aaccaatttgaattgttcattccgagttattaataatcaaatgttaataacttaacga 1020

QY 1021 AAGAAGTGA 1029  
|||||||

Db 1021 aagaagtga 1029

RESULT	5
X60830	
ID	X60830 standard; DNA; 3336 BP.

AC X60830;

DE S. aureus riboflavin biosynthesis operon.

KM Lumazine synthase; ribH; riboflavin; differential screening; gastritis  
KM H. pylori infection; gastrointestinal carcinoma; gastric ulcer; wound;  
KM bacterial adhesion; cell invasion; riboflavin biosynthesis operon; ss.

OS Staphylococcus aureus.

PN W09927128-A1

PD 03-JUN-1999

PF 24-NOV-1998; 98WO-US25096.

PR 25-NOV-1997; 97US-0977553.

PA (SMIK ) SMITHKLINE BEECHAM CORP

PI Debouck C, Fedon JC, Jaworski DD, Mooney J, Palmer IM:

XX  
F1  
112111 CM, wang, walter RL, zhong L,

DK WFL; 1999-33/030/30  
XX

PT New Staphylococcus aureus ribh polynucleotide and polypeptide useful in the treatment of gastric ulcers and gastritis

PS Example 2; Page 34-36; 48pp; English.

The invention relates to polynucleotides and polypeptides of the *lumazine synthase* (beta -subunit) family, designated *ribH*. They are isolated from *Staphylococcus aureus*. *ribH* enzymes are involved in the synthesis of *riboflavin*. The polynucleotides and polypeptides may be employed as research reagents and material for the discovery of treatments and diagnostics for diseases, particularly human diseases. They can be used for diagnostics of the disease and staging of disease, and as reagents in differential screening methods. The polynucleotides may be used as a source for hybridization probes, and for screening of genetic mutations, serotype, and identification, and for chromosome identification. The polypeptides can be used to produce antibodies. The polypeptides can also be used in vaccine formulations, and to identify agonists and antagonists. These are used to prevent, inhibit or treat diseases, particularly of *Helicobacter pylori* infections, such as gastrointestinal carcinoma. Gastric ulcers and gastritis may also be treated. The polypeptides can also be used to treat wounds and in-dwelling devices to prevent bacterial adhesion and infection, and to block *ribH* protein-mediated mammalian cell invasion. Antagonists and agonists of *ribH* may also be employed to inhibit and treat diseases. The present sequence represents the *S. aureus riboH* gene. The present sequence encodes genes *ribG*, *ribA* and *ribH*.

**SQ** Sequence 3336 BP; 1202 A; 459 C; 678 G; 997 T; 0 other;

Query Match	100.0%	Score 1029;	DB 20;	Length 3336;
Best Local Similarity	100.0%;	Pred. No. 7.6e-230;		
Matches 1029; Conservative	0;	Mismatches	0;	Gaps 0;
		Indels	0;	

QY 1 ATGGATTATGCGATTCAACTTGCAATATGCTACAAGGTCAAACAGGTGTTAATCCACCC 60

Db	1	atgattatcgattcgaattccaattgcacaatatgatacaagttcaaacaggtgttaattccacc	60
Qy	61	gTTCGGCCGTGTGTAGATTAAAGTAGAGATTGTTGGTATTTGGTGCACACTTGGAGAAA	120
Db	61		120
Db	61	gttgccgcgtctgtaagtaataagaagtaagattgtctgtgatttgytgacaacttggaaaaa	120
Qy	121	GGTACACAGATGGCGAGCTTCAAGCAAGCACTTGATATGGCAACAAATAATGCTGAAGTGG	180
Db	121	ggtgacaagaatcgsgsggttcaagaacctgatgatgagcaacaacaaaaatgctgaaggtgsg	180
Qy	181	ACGATTTATATTTACGTTTAGAGCCATGTAGTCATTTTGGTTCAACACCACCGTGTGTTAC	240
Db	181		240
Db	181	acgattatattactgattagagccattgtagcatattgtgttcaaacaccacccgtgttaac	240
Qy	241	AAATTTATTTGATTATAGATAGCAAAAAGTGTATACGCAACAAAAGCAATTCGTTAGAC	300
Db	241		300
Db	241	aaaatttctgattgtaagaatagcataaagtctatacgcacaacaaagacattcgttagac	300
Qy	301	ACACATGCTGATGAGACGTTACGGGCTCACGGTATTTGAGGTTGCAATGCGTTGATGATGAA	360
Db	301		360
Db	301	acacatgtgattgtagagcgtttagcggtctcaacgatttgaagttgaatgctgtatgata	360
Qy	361	CGGGCATCACAAATTTATTCGCAAGCAACTTTTTTAAAGCAAAAAGCAAGCAACTGCCACAAAT	420
Db	361		420
Db	361	cggggatccacaattataccaaagactttttaaagcaaaagcaaaagcaaatgcacaatat	420
Qy	421	ACAATGAAAGTATCTGCAACTTTAGATGTGTTAAACAAGCCANATGTATATGAGCAAAATCMA	480
Db	421		480
Db	421	acagttgaagatctctgtaagtttagatgtgtaaaccaagcgaatgataigtgacaagttcaa	480
Qy	481	TGGATTTACTTAACAAAGAGGTTAAACAAGATGCTATATAAGTTAAAGACATGACACGACGA	540
Db	481		540
Db	481	tggattacttaacaagaaggttlaacaagaatgctataaagttlaagaaacttgcacaagcga	540
Qy	541	GTTGTTACTGGAAGACCTAGACTTGAATTTGATGATGCCAATATATCTACAGCTATTTGCA	600
Db	541		600
Db	541	gtgttaactcgtgaagaagcgttaacgttgaatgataatgataccacaatactataccatgattcaa	600
Qy	601	GATGGAAGAAAAACCTATATAAAGTAAATATTTGCTCAAGTCTGGGAATATTCATTTTAAATCAG	660
Db	601		660
Db	601	gatggaaaaaaacctataaaaagtaataatgtctgaagttcgggaatatcatcttaactcag	660
Qy	661	CAATTTATCAAGATGATCAACCACAATTTGATATATACGAAATTCGAATTTTAAACA	720
Db	661		720
Db	661	caaatcttacaagatgataccaacaacttggatataatacgtgaaaaatccaaattlaaaca	720
Qy	721	AGCAATCAAAACACATATTTGAATTTATTTACTTGAATGCTTTGGATTTAAACAATTTCTT	780
Db	721		780
Db	721	agcatccaacacatatgtaaatattattacttgaagttcttgatttaacaacaacttctt	780
Qy	781	CACATTTTATATAAAGAGAGATTTGCAACTTTCCTAGTGTGAGGACAGTCCAAACCACTACT	840
Db	781		840
Db	781	cacaatttataataaagagatgttgaacttgcctagtcgaaggtccaaacacactact	840
Qy	841	TCAGAATTCACATCTTTATATAGATGAATTTATCTGATTTATTCGCCGAATTTAAT	900
Db	841		900
Db	841	tcgaattcttccaatcataataataagatgaattatctctatattatgccccgaattaat	900
Qy	901	GGCGGATCTGGAATTTATTCATTTTATCAACAATGATGTGATTGAGATTACACAGATGCG	960
Db	901		960
Db	901	ggcggatctggaatttcaatttcaatttcaacaacaatgaltgattgataaccagatgsg	960
Qy	961	AACCAATTTGAAATTTGTTCAATTCGCCGATTTATTAATCAAAAAGTTAAATTAATCTTACGA	1020
Db	961		1020
Db	961	aaccaacttggaaatgtctcatctccgagttattaaatccaanaatgtaataacttaacttaca	1020
Qy	1021	AAGAAGTGA 1029	
Db	1021		
Db	1021	aagaaagtga 1029	

XX	V74453/c
ID	V74453 standard; DNA; 7588 BP.
AC	
XX	V74453:
XX	
DT	16-MAR-1999 (first entry)
DE	Staphylococcus aureus contig SBQ ID #142.
XX	
KW	Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW	cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW	skin infection; surgical wound infection; scalded skin syndrome;
RX	toxic shock syndrome; ds.
XX	
OS	Staphylococcus aureus.
XX	
FH	Key Location/Qualifiers
FT	misc_feature 661..720
FT	/tag= a
FT	/note= "these bases represent a line of missing text in
FT	the sequence listing in the specification. They
FT	are included to maintain the nucleotide numbering
FT	given in the specification for this DNA sequence"
FT	misc_feature 2461..2520
FT	/tag= b
FT	/note= "these bases represent a line of missing text in
FT	the sequence listing in the specification. They
FT	are included to maintain the nucleotide numbering
FT	given in the specification for this DNA sequence"
FT	misc_feature 4261..4320
FT	/tag= c
FT	/note= "these bases represent a line of missing text in
FT	the sequence listing in the specification. They
FT	are included to maintain the nucleotide numbering
FT	given in the specification for this DNA sequence"
FT	misc_feature 6061..6120
FT	/tag= d
FT	/note= "these bases represent a line of missing text in
FT	the sequence listing in the specification. They
FT	are included to maintain the nucleotide numbering
FT	given in the specification for this DNA sequence"
PN	EP786519-A2.
XX	
PD	30-JUL-1997.
XX	
PF	07-JAN-1997: 97EP-0100117.
PR	05-JAN-1996: 96US-0009861.
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
P1	Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;
P1	Rosen CA;
DR	WPI: 1997-374922/35.
XX	
PT	Polynucleotide(s) and proteins derived from Staphylococcus aureus
PT	stored on computer readable medium and used in the production of
XX	anti-S.aureus vaccines
PS	Claim 1; Page 740-744; 3271pp; English.
CC	
CC	This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC	of the invention. The DNA sequences are recorded on a computer readable
CC	medium, preferably selected from a floppy or hard disk, random access
CC	memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC	the S.aureus DNA sequences allows putative functions to be assigned so
CC	that protein-encoding or regulatory regions of commercial, therapeutic or
CC	industrial importance can be obtained. Specifically, sequences which are
CC	likely to encode antigens have been identified and these polypeptides can
CC	be used in a vaccine composition against S.aureus infection. The
CC	polypeptides can also be used in a kit for the immunodetection of



CC S. aureus in a sample. S. aureus is implicated in numerous human diseases, including cellulitis, eyelid infections, food poisoning, osteomyelitis, skin and surgical wound infections, scalded skin syndrome, toxic shock syndrome, etc. Organisms transformed with the DNA sequences can be used for recombinant production of the polypeptides. The new DNA sequences (and their fragments) are useful as primers or probes for isolating CC homologues of any of the S. aureus DNA sequences contained on the CC computer readable medium.

SQ Sequence 7588 BP; 2333 A; 1341 C; 1061 G; 2613 T; 240 other;

Query Match 97.2%; Score 1000; DB 18; Length 7588;

Best Local Similarity 98.7%; Pred. No. 5.1e-223;

Matches 1019; Conservative 0; Mismatches 10; Indels 3; Gaps 1;

```

QY 1 ATGGATTATGCGATTCACATTCGCAATATGTGTACAAAGTCCAAAGAGTGAATCCACC 60
DB 5849 ATGGATTATGCGATTCACATTCGCAATATGTGTACAAAGAGTGAATCCACC 5790
QY 61 GTTGGCCCTGTGTAGTAAATGAAGTAGAGATTGTGTATGTGTGACACACTTGAGAAA 120
DB 5789 GTTGGCCCTGTGTAGTAAATGAAGTAGAGATTGTGTATGTGTGACACACTTGAGAAA 5730
QY 121 GGTGACACAGATGCGGAGGTTCAAGCACTTGATATGCGACACAAATGCTGAAGGTGCG 180
DB 5729 GGTGACACAGATGCGGAGGTTCAAGCACTTGATATGCGACACAAATGCTGAAGGTGCG 5670
QY 181 ACGATTATATTCAGTTAGAGCCATGTATCATTTTGGTTCAACACCACTTGCTTAAAC 240
DB 5669 ACGATTATATTCAGTTAGAGCCATGTATCATTTTGGTTCAACACCACTTGCTTAAAC 5610
QY 241 AAAATTATGATTGTAGATAGCAAAAAGTATATACGCAACAAAGAGCAATTCCTTAGAC 300
DB 5609 AAAATTATGATTGTAGATAGCAAAAAGTATATACGCAACAAAGAGCAATTCCTTAGAC 5550
QY 301 ACGATGCGTATGAGAGATTACGGGCTCAGGGTATGAGGTTGAATCGTTGATGATGAA 360
DB 5549 ACGATGCGTATGAGAGATTACGGGCTCAGGGTATGAGGTTGAATCGTTGATGATGAA 5490
QY 361 CGGCGATCACAATATATACCAAGACTTTTAAAGCAAAAGCAAGCGCCCAATTT 420
DB 5489 CGGCGATCACAATATATACCAAGACTTTTAAAGCAAAAGCAAGCGCCCAATTT 5430
QY 421 ACAAGTGAAGTATCTGCAAGTTTATGATGTAACAAAGCAAGTATGATGCAAGTCAA 480
DB 5429 ACAAGTGAAGTATCTGCAAGTTTATGATGTAACAAAGCAAGTATGATGCAAGTCAA 5370
QY 481 TCGATTACTTAACAAGAGGTTAAACAAGATGCTATTAAGTTAAGACATCGACGACGCA 540
DB 5369 TCGATTACTTAACAAGAGGTTAAACAAGATGCTATTAAGTTAAGACATCGACGACGCA 5310
QY 541 GTGTACTGGAAGAGCTACAGTTGAAATGATGATCCACAATTTACTACCGTATTCAA 600
DB 5309 GTGTACTGGAAGAGCTACAGTTGAAATGATGATCCACAATTTACTACCGTATTCAA 5250
QY 601 GATGGAAGAAACCCCTATAAAGTATATGCTCTAAGTCTGGGAATATTCATTTTAAATCAG 660
DB 5249 GATGGAAGAAACCCCTATAAAGTATATGCTCTAAGTCTGGGAATATTCATTTTAAATCAG 5190
QY 661 CAAATTTATCAAGTATCAACACCAATTTGGATATATACCTGAAATTCCAAATTTTAAACA 720
DB 5189 CAAATTTATCAAGTATCAACACCAATTTGGATATATACCTGAAATTCCAAATTTTAAACA 5130
QY 721 AGCAATCAACACATATATGAATATTTACTCTGAGTCTGTGTGAATTAACAACAATTCCTT 780
DB 5129 AGCAATCAACACATATATGAATATTTACTCTGAGTCTGTGTGAATTAACAACAATTCCTT 5070
QY 781 CACAATTTATATAAAGAGAGTGTGAATTTGCTAGCTGAGGAGGCTCCAAACCACTACT 840
DB 5069 CACAATTTATATAAAGAGAGTGTGAATTTGCTAGCTGAGGAGGCTCCAAACCACTACT 5010
QY 841 TCAGAAATTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 897

```

```

DB 5009 TCAGAAATTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
QY 898 ATGGCGGATCGGAATTTATCATTTTATCAAAATGATGATGATGATGATGATGATGATGATGAT
DB 4949 ATGGCGGATCGGAATTTATCATTTTATCAAAATGATGATGATGATGATGATGATGATGATGAT
QY 958 GCGAACCATTGGAATTTGTCATTCGAGTTATTAATCAAAATGTTAAATTAATTAATTAATTAAT
DB 4889 GCGAACCATTGGAATTTGTCATTCGAGTTATTAATCAAAATGTTAAATTAATTAATTAATTAAT
QY 1018 CGAAAGAGTGA 1029
DB 4829 CGAAAGAGTGA 4818

RESULT 7
T84023
ID T84023 standard; DNA; 1269 BP.
XX
AC T84023;
XX
DT 01-SEP-1998 (first entry)
XX
DE DNA encoding a Riboflavin biosynthesis protein RIBG.
XX
KW Staphylococcus aureus protein; ribozyme; antisense sequence; control;
KW Staphylococcal gene; regulatory element; bacterial gene expression;
KW vaccine; Staphylococcal infection; food poisoning; scalded skin syndrome;
KW toxic shock syndrome; ss.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT CDS 734..1033
FT /tag= a
XX
XX WO9730070-A1.
XX
XX 21-AUG-1997.
XX
XX 19-FEB-1997; 97WO-US02318.
XX
XX 20-FEB-1996; 96US-0011888.
XX
XX (SMIK ) SMITHKLINE BEECHAM CORP.
XX
XX Black MT, Burnham MK, Hodgson JE, Knowles DJC, Nicholas RO;
XX Pratt JM, Reichard RW, Rosenberg M, Ward JM;
XX WPI; 1997-424969/39.
XX P-PDB; W28084.
XX
XX Novel polypeptide(s) from Staphylococcus aureus strain WCUH29 - used
XX to isolate antimicrobial compounds, and in vaccines against S.
XX aureus infection
XX
XX Claim 9; Page 825; 989pp; English.
XX
XX The present sequence encodes a Staphylococcus aureus protein, that,
XX based on homology with an Bacillus subtilis protein, is believed to be a
XX Riboflavin biosynthesis protein RIBG. The present sequence was obtained
XX from a library of clones of S. aureus WCUH 29 in Escherichia coli. The
XX DNA sequence can be used in the construction of ribozymes and antisense
XX sequences to control the expression of Staphylococcal genes. The DNA
XX sequence is also useful as a source of regulatory elements for the
XX control of bacterial gene expression. The encoded protein may be used
XX to produce vaccines to enable a host to produce specific antibodies
XX with antibacterial action. These vaccines and antibodies would protect
XX a host against invasion by S. aureus, and conditions relating to
XX Staphylococcal infection, e.g. Staphylococcal food poisoning, scalded
XX skin syndrome, and toxic shock syndrome.

```

PI Black MT, Burnham MKR, Fedon JC, Hodgson JE, Knowles DJC;  
PI Lonetto MA, Nicholas RO, Palmer LM, Pratt JM, Reichard RW;

AC Q10165,



```

DT 10-SEP-1999 (first entry)
XX
XX B. subtilis rib operon nucleotide sequence.
DE
XX Riboflavin: open reading frame; ORF: structural gene; promoter;
KW vitamin B2; Bacillus subtilis; rib operon; ds.
XX
OS Bacillus subtilis.
XX
XX Key Location/Qualifiers
FH 1..5567
FT CDS
FT
FT /*tag= a
FT /note= "reading frame 1 translated protein (Y21801)"
FT
FT CDS
FT
FT /*tag= b
FT /note= "reading frame 2 translated protein (Y21802)"
FT
FT CDS
FT
FT /*tag= c
FT /note= "reading frame 3 translated protein (Y21803)"
FT
XX
XX US5925538-A.
XX
XX 20-JUL-1999.
XX
XX 24-AUG-1998; 98US-0138775.
XX
XX 11-SEP-1990; 90US-0581048.
XX 22-JUN-1989; 89US-0370378.
XX 21-APR-1992; 92US-0873572.
XX 06-FEB-1995; 95US-0384626.
XX 24-AUG-1998; 98US-0138775.
XX
XX (HOFF ) ROCHE VITAMINS INC.
XX
XX Erdenberger T, Hatch RT, Hermann T, Perkins JB;
XX Pero JG, Sloima A;
XX
XX WPI: 1999-418271/35.
XX P-PSDB: Y21801, Y21802, Y21803.
XX
XX Use of bacterial strains for the over production of riboflavin
XX
XX Example 2: Flg 3A-S; 61pp; English.
XX
XX The invention relates to a method of producing riboflavin that comprises,
XX providing a recombinant bacterium containing an exogenously introduced
XX bacterial ORF (open reading frame) 3 (or ORF5) structural gene and an
XX exogenous promoter, culturing the bacterium, and recovery of the product.
XX The method is useful for the production of large quantities of riboflavin
XX (vitamin B2). The present sequence represents the complete nucleotide
XX sequence of B. subtilis rib operon (also see X81948 for the antisense
XX sequence).
XX
XX Sequence 5567 BP: 1751 A; 1096 C; 1297 G; 1423 T; 0 other;
SQ

```

```

Query Match 11.7%; Score 120.4; DB 20; Length 5567;
Best Local Similarity 51.0%; Pred. No. 7.6e-19;
Matches 317; Conservative 0; Mismatches 296; Indels 9; Gaps 1;

```

```

QY 1 ATGGATTTATGCGATTCACAAATATGATGACAAAGTCAACAGGTTTATCCACCC 60
DB 1116 atgaagctgacctagatctgcgaagcagcgagacacgcaatccatccgctc 1175
QY 61 GTTGGCGGTGTGATGATGAGGTAGATTGTTGATTTGGTGCACACTTGGAGAAA 120
DB 1176 gtccgcgcgtgtcgttaagaagcacaatctgcggaatg9gcgccatlttaaat 1235
QY 121 GGTGACACGATCGGAGGTTCACCACTTGTATGGCACAACAATAAGCTGAAGTCGC 180
DB 1236 ggtgaagcctcagcagaatcgtcatcgtatgctgagcagcagatcagaaggtg 1295
QY 181 ACGATTATATACGTAGAGCCATGTAGTCAITTTGGTTCAACACACCCTGTGTAAAC 240

```

```

DB 1296 gacattacgttacactcgcgaaccgctgcagccattacgaaacacccgcatgtcagaa 1355
QY 241 AAAATTATGATTTAGTATGACAAAGTAGTATACGCAACAAAGG-----CAAT 291
DB 1356 ttgattalcaactctggtatcaaaagagctcgtg9cagatgaagatcctaaccgctt 1415
QY 292 TCGTTAGACACATGATGATGATGACGTTACGGGCTCACGCTATTGAGTTGATGCGTT 351
DB 1416 gtg9cgtggaagagagatcagcatgatgaagaagctgcatgtgaagtaaggaagcctc 1475
QY 352 GATGATGAACGGGCATCACATTTATACCAAGCTTTTAAAGCAACAAAGCAACATCG 411
DB 1476 ctgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1535
QY 412 CCACAAATTCAGATGCAAGTATCTGCAAGTATGATGATGATGATGATGATGATGATG 471
DB 1536 ccgtacgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1595
QY 472 CAAAGTCAATGATTTACTATACCAAGAGGTTAAACAGATGCTATATAGTTAAGATCGA 531
DB 1596 gacagcaaatgcatcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1655
QY 532 CACGACGCGATGTTACTGGAAGACGTACAGTGTGATTTGATGATGATGATGATGATG 591
DB 1656 caccgaagcatttgcgtgagcagcagcagcagcagcagcagcagcagcagcagcagc 1715
QY 592 CGTATTCAAGATGATAAAGAC 613
DB 1716 agactgcggaatgtatacaaaac 1737

```

```

RESULT 11
X81948/c
ID X81948 standard; DNA: 5567 BP.
XX
XX X81948;
XX
XX 10-SEP-1999 (first entry)
XX
XX B. subtilis rib operon complementary nucleotide sequence.
XX
XX Riboflavin: open reading frame; ORF: structural gene; promoter;
XX vitamin B2; Bacillus subtilis; rib operon; complementary; ds.
XX
XX Bacillus subtilis.
XX
XX Key Location/Qualifiers
FH 1..5567
FT CDS
FT
FT /*tag= a
FT /note= "reading frame 1 translated protein (Y21804)"
FT
FT CDS
FT
FT /*tag= b
FT /note= "reading frame 2 translated protein (Y21805)"
FT
FT CDS
FT
FT /*tag= c
FT /note= "reading frame 3 translated protein (Y21806)"
FT
XX
XX US5925538-A.
XX
XX 20-JUL-1999.
XX
XX 24-AUG-1998; 98US-0138775.
XX
XX 11-SEP-1990; 90US-0581048.
XX 22-JUN-1989; 89US-0370378.
XX 21-APR-1992; 92US-0873572.
XX 06-FEB-1995; 95US-0384626.
XX 24-AUG-1998; 98US-0138775.
XX
XX (HOFF ) ROCHE VITAMINS INC.
XX
XX Erdenberger T, Hatch RT, Hermann T, Perkins JB;
XX

```



BR 22-JUN-1989; 89US-0370378.  
 PR 21-APR-1992; 92US-0873572.  
 PR 06-FEB-1995; 95US-0384626.  
 XX  
 XX (HOFF ) HOFFMANN LA ROCHE INC.  
 PI Perkins JB, Pero JC, Sloma A;  
 DR WPI; 1999-023456/02.  
 DR P-PSDB; W81549-55.  
 XX  
 PT Nucleic acid encoding proteins involved in bacterial synthesis of  
 riboflavin - also mutant or recombinant bacteria producing  
 riboflavin at high levels  
 XX  
 PS Example 7; Fig 3A-S; 62pp; English.

CC This is the nucleotide sequence of the rib operon of *Bacillus*  
 CC *subtilis*. The rib operon includes an open reading frame (ORF)  
 CC encoding the beta-subunit (see W81553) of riboflavin synthase, and  
 CC 6 other ORFs, of which ORF1 is a partial coding region encoding a  
 CC 170-amino acid polypeptide (see W81555), ORF2 encodes a putative  
 CC rib-specific reductase (see W81554), ORF3 encodes a putative GTP  
 CC cyclohydrolase (see W81552), ORF4 encodes the riboflavin synthase  
 CC alpha subunit (see W81551), ORF5 encodes a putative rib-specific  
 CC deaminase (see W81550) and ORF6 encodes a 38 kDa protein (see  
 CC W81549) of unknown function. The invention provides vectors and  
 CC recombinant bacteria for overproducing riboflavin, in which nucleic  
 CC acids overproducing riboflavin biosynthetic proteins are introduced  
 CC into the chromosome of the host organism at multiple sites or in  
 CC multiple copies. The rib operon, e.g. comprising ORFs 2-5 and  
 CC the beta-riboflavin synthase gene, is used to make such recombinant  
 CC bacteria. The rib operon can be deregulated by replacing its  
 CC control regions with sequences that allow constitutive or  
 CC unregulated expression. The recombinant bacteria produce at least  
 CC 10 g/l riboflavin.  
 XX  
 SQ Sequence 5567 BP; 1749 A; 1102 C; 1298 G; 1418 T; 0 other;

Query Match 11.7%; Score 120.4; DB 20; Length 5567;  
 Best Local Similarity 51.0%; Pred. No. 7.6e-19;  
 Matches 317; Conservative 0; Mismatches 296; Indels 9; Gaps 1;  
 QY 1 ATGATTTATCGATTCAACTTGCATAATATGTACAGGTCAACAGGTTAATCCACC 60  
 DB 1116 atgaagctgcttagcttagctggaagcagcgagcagcgaatccatccgcctc 1175  
 QY 61 GTTGGCGCTGTGTATTAATGAAGTAGATTGTTGATGTCGACACTTGAGAAA 120  
 DB 1176 gtgcgctgctgtgtcgttaagacgaatgtcgtgaatgtgcgctcatttaaatat 1235  
 QY 121 GGTGACAGCATCGAGAGTTCAACACTTGTATATGCGACACAAATGCTGAGGTGG 180  
 DB 1236 ggtgaagctcagcagaagatcattcattcattcgtcgtgagcagatcagaaggtcgc 1295  
 QY 181 ACGATTATATTAGTTAGAGCCATGAGCTATTTGTTTCAACACACACCTGTGTTAAC 240  
 DB 1296 gacattcattcactcgaacgcgtcagccattcaggaacacgcgcatgtgcagaa 1355  
 QY 241 AAAATTATTGATTGATAGATACCAAAAGTAGTATACGCCAACAAAGA-----CAAT 291  
 DB 1356 ttgattatcaactcgtgtatcaaaagagtgctcgtgagatgagatcttaatccgctt 1415  
 QY 292 TCGTTAGACACATGATGATGACAGGTTACGGCTCAGCGATTGAGGTGATGCGTT 351  
 DB 1416 gtgcgtggaagagggatcagcagatgaaagaagctgagcatgaggttaaggaagcattc 1475  
 QY 352 GATGATGAACGGGATCACAAATTATACCAAGACTTTTAAAGCAAAAGCAAGCAACTG 411  
 DB 1476 ctgagcagcagcgagagagtgatgaaattctcgtcattatgagcagcgctt 1535  
 QY 412 CCACAATTACAGTGAAGTATCTGCAAGTTAGTGTAAACAGCAATGATATATGA 471

DB 1536 ccgtacgtcagctaaagcggctgcgcgccttgacggcgaagatagctaccagcagcgtt 1595  
 QY 472 CAAAGTCATGATTGATTACAAAGAGTTAAACAAGATGCTGTATAGTTAGACATGCA 531  
 DB 1596 gacgaagaatgatacagtcagaagcgtcgaacagagatgctcagcaataacagaagaaca 1655  
 QY 532 CACGACGAGTGTAACTGGAAGAGCTACAGTTGATTAAGATGATCCACAAATATCTACA 591  
 DB 1656 caccgaagatttagctgaggttgacacagtgaaagccgacaatccagcttaaccctgc 1715  
 QY 592 CGTATTCAGATGGAATAAAC 613  
 DB 1716 agactgccgaatgtaacaaac 1737  
 RESULT 13  
 ID 293752 standard; DNM; 5567 BP.  
 XX  
 AC 293752;  
 XX  
 DT 16-AUG-2000 (first entry)  
 XX  
 DE Rib operon of *Bacillus subtilis*.  
 XX  
 KW Riboflavin; rib operon; yeast; bacteria; transformation;  
 KW transcription element; gene expression; hair loss; skin inflammation;  
 KW growth disorder; vision; vision; sight; ds.  
 XX  
 OS *Bacillus subtilis*.  
 XX  
 FH Key  
 FT CDS  
 FT Location/Qualifiers  
 FT 1..5567  
 FT /tag= a  
 FT /note= "Contains in frame stop codons"  
 FT  
 FT RBS  
 FT 62..68  
 FT /tag= b  
 FT  
 FT CDS  
 FT 363..680  
 FT /tag= c  
 FT /label= ORF 6  
 FT 708..748  
 FT /tag= d  
 FT /label= Rho independent transcription terminator  
 FT  
 FT promoter  
 FT 771..799  
 FT /tag= e  
 FT /label= p1 promoter  
 FT 1034..1067  
 FT /tag= f  
 FT /label= Rho independent transcription terminator  
 FT 1084..1095  
 FT /tag= g  
 FT /label= ORF5  
 FT 1422..1433  
 FT /tag= h  
 FT /label= ORF5  
 FT 1442..1453  
 FT /tag= i  
 FT /label= RBS  
 FT 2177..2194  
 FT /tag= j  
 FT /label= CDS  
 FT 2197..2844  
 FT /tag= k  
 FT /label= ORF4  
 FT 2528..2556  
 FT /tag= l  
 FT /label= p2 promoter  
 FT 2843..2854  
 FT /tag= m  
 FT /label= CDS  
 FT 2859..4055  
 FT /tag= n  
 FT /label= ORF3  
 FT 4071..4085  
 FT /tag= o  
 FT /label= CDS  
 FT 4088..4552

```

FT      /tag- p
FT      /product= Beta riboflavin synthase
FT      4545..4574
FT      promoter
FT      /tag- g
FT      /label= p3 promoter
FT      4653..4659
FT      RBS
FT      /tag- t
FT      4665..5039
FT      CDS
FT      /tag- s
FT      /label= ORF2
FT      5038..5090
FT      terminator
FT      /tag- t
FT      /label= Rho independent transcription terminator
FT      complement (5053..5567)
FT      /tag- u
FT      /label= ORF1
FT      CDS
FT      EP1001026-A2.
XX      17-MAY-2000.
XX      PD
XX      PF
XX      22-JUN-1990; 99EP-0124888.
XX      PR
XX      22-JUN-1989; 89US-0370378.
XX      PR
XX      22-JUN-1990; 90EP-0111916.
XX      PA
XX      (HOFF ) HOFFMANN LA ROCHE & CO AG F.
XX      PI
XX      Perkins JB, Pero JG, Sloma A:
XX      WPI, 2000-320439/28.
XX      DR
XX      P-PSDB; Y83269, Y83270, Y83271.
XX      PT
XX      Novel vector, used for the large scale production of riboflavin,
XX      PT      comprises a bacterial or yeast riboflavin biosynthetic protein under
XX      PT      the control of a foreign transcription element
XX      PS
XX      Claim 1; Figure 3; 66pp; English.
XX      CC
XX      Vectors comprising a nucleic acid sequence of bacterial or
XX      CC      yeast origin, coding for one or more riboflavin biosynthetic proteins,
XX      CC      and one or more transcription elements not naturally associated with
XX      CC      the nucleic acid sequence are new. The vector and bacteria transformed
XX      CC      with it are used for the large scale production (over 10g/l) of
XX      CC      riboflavin. The riboflavin produced can be used to treat disorders
XX      CC      associated with riboflavin deficiency such as hair loss, skin
XX      CC      inflammation, vision deterioration and growth failure.
XX      CC
XX      Sequence 5567 BP; 1751 A; 1097 C; 1296 G; 1423 T; 0 other;
SQ
Query Match      11.7%; Score 120.4; DB 21; Length 5567;
Best Local Similarity 51.0%; Pred. No. 7.6e-19;
Matches 317; Conservative 0; Mismatches 296; Indels 9; Gaps 1;
QY      1 ATGCAATTATGCGATTCAACTTGCACAAATATGTCACAAAGTCAGGTGTTATCCACCC 60
DB      1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY      1116 atgaagctggccttagactctgcgaagcagggcgaagagacagccatccatccgctc 1175
DB      1116 gtgcggcgtctgtcgttaaaagcagcaaatctgcgaaatgagcgccattaaatat 1235
QY      61 GTTGGCGCTTGTAGTATTAAGTGAAGTATGTTGGTATGTCACACTTGAGAAAA 120
DB      1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY      1176 gtgcggcgtctgtcgttaaaagcagcaaatctgcgaaatgagcgccattaaatat 1235
DB      121 GGTGACAGCATGCGAGGTTCAGCACTGATATGTCACACAAATATGCTGAAGTGGC 180
QY      1236 ggtgaagctatgcaggaagctatgcatacatatgctggaagacatgcaagaggtgcc 1295
DB      181 ACATTTATATTAAGTGAAGCATGTATGTCATTTGGTTCAACACCCACCCCTGTGTAAAC 240
QY      1296 gacattacgttaacacgcgacccgtgcagcatctaacgaaacacccgcatgtgcagaa 1355
DB      241 AAATTTATGATTTGTAAAGTAGCAAAAGTGTATAGCCACCAAAAGA-----CAAT 291
QY      11111 1111 1111 1111 1111 1111 1111 1111 1111 1111

```

```

DB      1356 ttgatctaacctctggtatcaaaaagtgctggtgcgatgagagatcctaaccgctt 1415
QY      292 TCGTTAGACACACATGCTGATGAGACCGCTTACGGGCTACGGTATGTAGTGAATGCGTT 351
DB      1416 gtgcgtcggaaagagagatcaagcatgatgaagaagcctgcatctgaggggaaagagcatc 1475
QY      352 GATGATGAACGGCAGTCACAAATTAATACCAAGACTTTTAAAGCAAAAGCAACCAACTG 411
DB      1476 ctgcagacaccagcgaggaagcctgatcatgaaatcttcgactttatgagacagcgctt 1535
QY      412 CCACAAATTAACAGTGAAGTATCTGCAAGTTTATGATGTAACAGCAATGATTAATGA 471
DB      1536 ccgtacgtccacgcgtcaaaagcgcgtgcacgctctgcgcgaagatagctaccagacgggt 1595
QY      472 CAAAGTCATGATGATTACTTAACAAGAGTAAACACAAATGCTATTAAGTAAACATCGA 531
DB      1596 gacagcaaatggtatcagtcagacagcgcgcaagaagatgctcagcaataacagaaaca 1655
QY      532 CACGACGCGAGTGTAACTGGAAGACGTACAGTTGAATTAGATCCACATATATCTACA 591
DB      1656 caccaaagcattttagtcgagtggtgcacagtggaagccgacacatccgagcttaacctgc 1715
QY      592 CGTATTCAGATGGAATAAAC 613
DB      1716 agactgccgaatgtlaacaaaac 1737

RESULT 14
V52173/c
ID      V52173 standard; DNA; 14273 BP.
XX
XX      V52173;
XX
DE      23-OCT-1998 (first entry)
XX
XX      Streptococcus pneumoniae genome fragment SEQ ID NO:40.
XX
XX      Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
XX      KM      computer readable medium; vaccine; pharmaceutical composition; ds.
XX      OS
XX      Streptococcus pneumoniae.
XX
XX      W09818931-A2.
XX
XX      07-MAY-1998.
XX
XX      30-OCT-1997; 97WO-US19588.
XX
XX      31-OCT-1996; 96US-0029960.
XX
XX      (HUMA-) HUMAN GENOME SCI INC.
XX
XX      Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M:
XX      Kunsch CA, Rosen CA;
XX      WPI; 1998-272225/24.
XX
XX      Computer-readable medium with recorded Streptococcus pneumoniae
XX      PT      polynucleotide sequences - useful in diagnostic kits and assays, and
XX      PT      pharmaceutical compositions and vaccines for Streptococcus
XX      PT      pneumoniae
XX
XX      Claim 1; Page 389-398; 1409pp; English.
XX
XX      The present invention describes a computer readable medium which has
XX      CC      the nucleotide sequences SEQ ID NO:1 to 391 (V52134 to V52524) recorded
XX      CC      on it, or a representative fragment or a sequence at least 95% identical
XX      CC      to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in SEQ ID NO:1
XX      CC      to 391 (V52134 to V52524) are genomic fragments from Streptococcus
XX      CC      pneumoniae. The present invention also describes an isolated nucleic acid
XX      CC      molecule encoding a homologue of any of the fragments of the S. pneumoniae
XX      CC      genome (SEQ ID NO:1 to 391) where the nucleic acid molecule is produced
XX      CC      by a process comprising: (a) screening a genomic DNA library using as a

```

CC probe at target sequence defined by any of the sequences in SEQ ID NO:1  
CC to 391, identifying members of the library which contain sequences  
CC that hybridise to the target sequence and isolating the nucleic acid  
CC molecules from the members; or (b) isolating mRNA, DNA or cDNA produced  
CC from an organism, amplifying nucleic acid molecules whose nucleotide  
CC sequence is homologous to amplification primers derived from the  
CC fragment of the *S. pneumoniae* genome to prime the amplification and  
CC isolating the amplified sequences. The computer readable medium can be  
CC used in a computer-based system for identifying fragments of the  
CC *S. pneumoniae* genome of commercial importance, or expression modulating  
CC fragments of the *S. pneumoniae* genome. Products from the present  
CC invention can be used in diagnosis kits and assays, and pharmaceutical  
CC compositions and vaccines for *S. pneumoniae*.  
CC

SQ Sequence 14273 BP; 4357 A; 2428 C; 2694 G; 4794 T; 0 other;

Query Match	11.1%	Score	114.2	DB	19	Length	14273
Best Local Similarity	49.3%	Pred. No.	2.7e-17				
Matches	387	Conservative	0	Mismatches	388	Indels	10
						Gaps	3

Oy	1	ATGGATTATGCACTTCACTTCACATATATGATACAGCTCAACAGGGTTATATCCACC	60
Db	10928	ATGAAATTAGCAATTAACCTGGCACAAAAAGGGCGTGGTTACGTCAATTCCTATG	108659
Oy	61	CTTGGCGCTGTGTGATTAATGAAAGGTAGATTGTGGTATTTGGTCACTTGGAAAA	120
Db	10868	GTCCGCGCAGTATTGTATTAAGATATATCAATATTATCGACAAAGGTATTCATGATTTT	10809
Oy	121	GGTGACAGCATGGGAGGTTCAACCACTGATGTATGGCACAAACAAATGCTGAAGTGG	180
Db	10808	GGTGGCCCACTGCTGAGAGAAATCTCTTTAAAACTGTAGAAATCCCTGTGGAGCG	10749
Oy	181	ACGATTATTATTTACGTTAGAGCATGTAGTCATTTTGGTCAACCAACCCCTGTATAC	240
Db	10748	ACGCTTTATGTAACTGTCAACCTCGTGTGTCTACTTCGGGAAAAACCTCCGTATAGAT	10689
Oy	241	AAATTTATTTGATGTGAAGTATACCAAAAGTACTATATACGCACAAAAA-----CAATT	292
Db	10688	GCTATTAATCGATAGGTGATTATCAAGAGTACGATTGGAAGGCTGACTGTATCTTAT	10629
Oy	293	CGTTAGACACACATGATGATGAGAGCTTACCGGGCTCACGGTATTTGAGTTGAATCGTTG	352
Db	10628	GTAATCGAAAAAGAGTAAAGTACTTGGAGAAATATATCTTCAAGTTACTGTGGATTT	10569
Oy	333	ATGATGAACGGGCATCACATTTATACCAAGACTTTTAAAGCA-AAAGCAAAAGCACTG	411
Db	10568	TTAGAAATATGAGTCTTAACTTATAAAAAGTTTACAAAGTATATTAACCCAGCATGTA	10509
Oy	412	CCACAATTTACGTAAAGTATCTCGCAAGTTAGATGTGAAGAAACAGCGCAATGATATGTA	471
Db	10508	CCCTATGTTTTTATTAATATGCAATGTTCMATGATGAGAAATATGCCCTAAACAAAT	10449
Oy	472	CAAACTCAATGATTTACTTAACAAAGAGGTTAAACAGAATGCTATATAGTTAAGACATGTA	531
Db	10448	CAATCAAAATGATTTACTGGAAGAAAGAAAGAAAGCAATGTCATCAGTACGACACCAT	10389
Oy	532	CACGACGAGTGTAACTGGAAGACGTACAGTGTGAATTAGATGATCCACAAATATCTACA	591
Db	10388	GTTACTGCAATTTATGCTGGGAGTCAATCTACTGTTATTTCAAGAGATCTTTTGGTCAGATGT	10329
Oy	592	CGTATTTCAAGATGAAAAACCCCTTAAAGATATATGTCTTAAAGTCTGGGAAATTTTCAAT	651
Db	10328	AGATTGGAGGAAGAAAAATCTCTATCCGTATCATATGCGATACACATTTTACGAATCTCT	10269
Oy	652	TTTAA-ATCAGCAAAATTTATCAGATGAATCAACACCAATTTGGATATATCTAGAAATTC	710
Db	10268	CTTACTCTAAATATGTTAAAAACACCAATATGATTTTAAAACTTACATGGCCACTTCCTCT	10209
Oy	711	AAATTTAACAGCATTCAAACCATATTTGAATTTTACTGTGAAGTCTGTGATTTTAACT	770
Db	10208	GAAAGCAAAATATAATGAAAGCTATATCAAAATCATGCTGTGAAATCTTTCATTAAG	10149

QY	771	AACAA	775
Db	10148	AAAAA	10144

## RESULT 15

ID	standard; DNA; 1101 BP.
1	

AC X77179;

DT 04-AUG-1999 (first entry)  
XX

S. pneumoniae ribg polypeptide encoding DNA.

KW Streptococcus pneumoniae; pyrimidine deaminase; pyrimidine reductase;

KW gastrointestinal carcinoma; gastric ulcer; gastritis; ss.

Streptococcus pneumoniae.

PN W09927126-A1.

PD 03-JUN-1999.

PF 23-NOV-1998;  
yy

PR 25-NOV-1997; 97US-0979616.  
YY

PA (SMIK ) SMITHKLINE BEECHAM CORP.  
XX

PI Black MT, Debouck C, Fedon JC, Knowles DT, Kosmatka AI, Tonet

PI Palmer LM, Shilling LK, Stodola RK, Wang M, Warren RL;  
PI Zhong Y.

XX  
DR  
WPT : 1999-347726/29

DR P-PSDB; Y17949.  
XX

PT New Streptococcus pneumoniae ribb polypeptide and polynucleotide useful in the treatment of gastric ulcer and gastritis

XX  
PS  
Claim 6: Page 5: 48pp: English

XX The invention relates to streptococcus pneumoniae pyrimidine de

CC and pyrimidine reductase (ribG). *S. pneumoniae* ribG and its ant

CC to expression or activity of ribg can be determined by analysin

CC nucleic acid sequence encoding ribg or detecting the ribg polyn

CC a sample. ribg can also be used to identify antagonists or agon

CC an immunological response in an animal. Antimicrobial compounds (e.g. antibiotics and antagonists of ribG) especially broad-spectrum antibiotics

CC may be of use in the treatment of *Helicobacter pylori* infection. This should decrease the advent of *H. pylori*-induced cancers, such as

CC and gastritis. The present sequence represents a DNA encoding the

cc s. pneumoniae ribg polypeptide.  
xx

sequence 1101 BP; 392 A; 180 C; 216 G; 313 T; 0 other;

Query Match 10 98: Score 112 6: DB 20: Length 1101:

Best Local Similarity 49.2%; Pred. No. 3.2e-17;  
Matches 386: Conservative 0: Mismatches 389: Indels 10: Gaps 1

1 ATG GAT TAT GCG AT TCA A C T T T G C A A T T A T T G T A C A A G G T C A A A C A G G T G T T A A T T C C A C C C 60

Db 19 atgaattaccaataaaactaacacaaaagaaactaattacatcaatcccaatcctata 78

61 GTGGCGCTGTGTAAGTTAATGAAGTAGGATTGTTGGTATTGGTGCACACTTGAGAAA 120

Db 79 attgacacattatgtaaaagataatcacattatcgacaaattatcatgaatttt 138

50 Sequence 1101 BP; 392 A; 180 C; 216 G; 313 T; 0 other,

Query Match	10.9%	Score 112.6;	DB 20;	Length 1101;
Best Local Similarity	49.28;	Pred. No. 3.2e-17;		
Matches 386;	Conservative	0;	Mismatches 389;	Indels 10;
				Gaps 3;

Oy	1	ATGCAATTTCGATTCACCTTCGCAAAATTTGGTACAAAGTCGCAACAGGCTTTAATTCACCC	60
Db	19	atgaattatgaacataaacaacgycacaaaaaaggccggtttagcgtcaacccaatccatcatg	78
Oy	61	GTTCGCGCTGTTGTAGTTAATGAAGTGAAGTATTTGGTATTTGGTCACCACTTGAAGAAA	120
Db	79	gttgcgcgaattatcgtataaagatcaatcaaccttcgagcaaggcttcatcgaatcttc	138



```
OY 121 GGTGACAGCATCGGAGTTCAAGCACTTGAATGGCACACAAAATGCTGAAGTGGC 180
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 139 ggtggccacatgctggagagaatgctcttaaaaacgtagaatatccctgctggagcg 198
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
OY 181 ACGATTATTATTAGTTAGAGCCATGATGATTTTGGTTCAACACCACCCCTGTGTAA 240
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 199 agccttaattgaacaccttgtaacctgtgtacctcggaataacacctcctgtatagat 258
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
OY 241 AAATTATTGATTGTAAGTAGCAAAAGTATACGACACAAAAGA-----CAATT 292
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 259 gctataatcgatagtgatatacaagagtgatcattggaagccctagacttaacccatt 318
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
OY 293 CGTTAGACACATGCTGATGAAGAGTTACGGGCTCACGGTATTGAGTTGAATGCTTG 352
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 319 gtatctggaaaagagaaagatattcttgagaataatctcaagtlactgttgatc 378
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
OY 353 ATGATGAACGGGATCATCATTTATACCAAGACCTTTTAAAGCA-AAAGCAAGCACTG 411
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 379 ttgaaaaatgagtgcttaacttaataaaaagtttagaaagtatattaccacgcatgta 438
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
OY 412 CCACAAATTACAGTGAAGATCTGCAAGTTAGATGTTAAACAAAGCAATGATATGA 471
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 439 ccctatgcttcttgaataatgcaatgcaatgcaatgcaatgcaatgcaatgcaatgcaat 498
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
OY 472 CAAAGTCAATGATTTACTTAACAAGAGGTTAAACAAAGATGCTATAAGTTAAGACATGA 531
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 499 caatccaatlgatctactgaagaagaagaagaagaagcatgcatcagttacgacacat 558
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
OY 532 CACGACGCACTGTTAACTGCAAGACGTTACGTTGAATTAGATGATCCACAATATACTACA 591
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 559 gtagtgcacattatggtggaagtcacatctgattcaagaagatcccttgcgtgacatg 618
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
OY 592 CGATTTCAGATGGAATAAAACCTATTAAGTAAAGTAAATTTGCTAAGTCTGGGAATATTCAT 651
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 619 agattggaggaaggaataatccatccgatacatatgcatatacaatctgaacccct 678
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
OY 652 TTTA-ATCAGCAAAATTTATCAAGATGAATCAACACCAATTTGGATATATCTGAAAATCC 710
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 679 cttaaccttaaaatcgtaaaaacagcaaatlgatataaacttaacatgcccactcct 738
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
OY 711 AAATTTAACAAGCAATCAACATATTTGAATTTATTTACTTGAAGTCTTGATTTAAC 770
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 739 gaagacaaaaataaatgaagctataatcaaatcatgctgtaaatattcattcataaag 798
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
OY 771 AACAA 775
    ||| |
Db 799 aaaaa 803
    ||| |
```

Search completed: May 7, 2001, 12:43:42  
Job time: 4383 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2001, 10:20:14 ; Search time 1037.25 Seconds  
(without alignments)  
8666.614 Million cell updates/sec

Title: US-09-376-633-1  
Perfect score: 1029  
Sequence: 1 ATGGATTATGCGATTCAACT.....TAACTTACGAAGAAGTGA 1029

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 9623517 segs, 4368049070 residues

Total number of hits satisfying chosen parameters: 19247034

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*  
12: gb\_est12:\*  
13: gb\_est13:\*  
14: gb\_est14:\*  
15: gb\_est15:\*  
16: gb\_est16:\*  
17: gb\_est17:\*  
18: gb\_est18:\*  
19: gb\_est19:\*  
20: gb\_est20:\*  
21: gb\_est21:\*  
22: gb\_est22:\*  
23: gb\_est23:\*  
24: gb\_est24:\*  
25: gb\_est25:\*  
26: gb\_est26:\*  
27: gb\_est27:\*  
28: gb\_est28:\*  
29: gb\_est29:\*  
30: gb\_est30:\*  
31: gb\_est31:\*  
32: gb\_est32:\*  
33: gb\_est33:\*  
34: gb\_est34:\*  
35: gb\_est35:\*  
36: gb\_est36:\*  
37: gb\_est37:\*  
38: gb\_est38:\*  
39: gb\_est39:\*  
40: gb\_est40:\*  
41: gb\_est41:\*  
42: gb\_est42:\*  
43: gb\_est43:\*

44: em\_esthum10:\*  
45: em\_esthum11:\*  
46: em\_esthum12:\*  
47: em\_esthum13:\*  
48: em\_esthum14:\*  
49: em\_esthum15:\*  
50: em\_esthum16:\*  
51: em\_esthum17:\*  
52: em\_esthum18:\*  
53: em\_esthum19:\*  
54: em\_esthum20:\*  
55: em\_esthum21:\*  
56: em\_esthum22:\*  
57: em\_esthum23:\*  
58: em\_esthum24:\*  
59: em\_esthum25:\*  
60: em\_esthum26:\*  
61: em\_esthum27:\*  
62: em\_esthum28:\*  
63: em\_esthum29:\*  
64: em\_esthum30:\*  
65: em\_esthum31:\*  
66: em\_esthum32:\*  
67: em\_esthum33:\*  
68: em\_esthum34:\*  
69: em\_esthum35:\*  
70: em\_esthum36:\*  
71: em\_esthum37:\*  
72: em\_esthum38:\*  
73: em\_esthum39:\*  
74: em\_esthum40:\*  
75: em\_esthum41:\*  
76: em\_esthum42:\*  
77: em\_esthum43:\*  
78: em\_esthum44:\*  
79: em\_esthum45:\*  
80: em\_esthum46:\*  
81: em\_esthum47:\*  
82: em\_esthum48:\*  
83: em\_esthum49:\*  
84: em\_esthum50:\*  
85: em\_esthum51:\*  
86: em\_esthum52:\*  
87: em\_esthum53:\*  
88: em\_esthum54:\*  
89: em\_esthum55:\*  
90: em\_esthum56:\*  
91: em\_esthum57:\*  
92: em\_esthum58:\*  
93: em\_esthum59:\*  
94: em\_esthum60:\*  
95: em\_esthum61:\*  
96: em\_esthum62:\*  
97: em\_esthum63:\*  
98: em\_esthum64:\*  
99: em\_esthum65:\*  
100: em\_esthum66:\*  
101: em\_esthum67:\*  
102: em\_esthum68:\*  
103: em\_esthum69:\*  
104: em\_esthum70:\*  
105: em\_esthum71:\*  
106: em\_esthum72:\*  
107: em\_esthum73:\*  
108: em\_esthum74:\*  
109: em\_esthum75:\*  
110: em\_esthum76:\*  
111: em\_esthum77:\*  
112: em\_esthum78:\*  
113: em\_esthum79:\*  
114: em\_esthum80:\*  
115: em\_esthum81:\*  
116: em\_esthum82:\*

117: gb\_est48:\*  
 118: gb\_est49:\*  
 119: gb\_est50:\*  
 120: gb\_est51:\*  
 121: gb\_est52:\*  
 122: gb\_est53:\*  
 123: gb\_est54:\*  
 124: gb\_est55:\*  
 125: gb\_est56:\*  
 126: gb\_est57:\*  
 127: gb\_est58:\*  
 128: gb\_est59:\*  
 129: gb\_est60:\*  
 130: gb\_est61:\*  
 131: gb\_est62:\*  
 132: gb\_est63:\*  
 133: gb\_est64:\*  
 134: gb\_est65:\*  
 135: gb\_est66:\*  
 136: gb\_est67:\*  
 137: gb\_est68:\*  
 138: gb\_est69:\*  
 139: gb\_est70:\*  
 140: gb\_est71:\*  
 141: gb\_est72:\*  
 142: gb\_est73:\*  
 143: gb\_est74:\*  
 144: gb\_est75:\*  
 145: gb\_est76:\*  
 146: gb\_est77:\*  
 147: gb\_est78:\*  
 148: gb\_est79:\*  
 149: gb\_est80:\*  
 150: gb\_est81:\*  
 151: gb\_est82:\*  
 152: gb\_est83:\*  
 153: gb\_est84:\*  
 154: gb\_est85:\*  
 155: gb\_est86:\*  
 156: gb\_est87:\*  
 157: gb\_est88:\*  
 158: gb\_est89:\*  
 159: gb\_est90:\*  
 160: gb\_est91:\*  
 161: gb\_est92:\*  
 162: gb\_est93:\*  
 163: gb\_est94:\*  
 164: gb\_est95:\*  
 165: gb\_est96:\*  
 166: gb\_est97:\*  
 167: gb\_est98:\*  
 168: gb\_est99:\*  
 169: gb\_est100:\*  
 170: gb\_est101:\*  
 171: gb\_est102:\*  
 172: gb\_est103:\*  
 173: gb\_est104:\*  
 174: gb\_est105:\*  
 175: gb\_est106:\*  
 176: gb\_est107:\*  
 177: gb\_est108:\*  
 178: gb\_est109:\*  
 179: gb\_est110:\*  
 180: gb\_est111:\*  
 181: gb\_est112:\*  
 182: gb\_est113:\*  
 183: gb\_est114:\*  
 184: gb\_est115:\*  
 185: gb\_est116:\*  
 186: gb\_est117:\*  
 187: gb\_est118:\*  
 188: gb\_est119:\*  
 189: gb\_est120:\*

Result	Score	Match	Length	ID	Description
1	70.4	6.8	1094	229	AL101513 Drosophila
2	68.2	6.6	420	30	AV426745 AV426745
3	63.2	6.1	1101	229	AL063921 Drosophila
4	63	6.1	469	141	BE920828 EST424597
5	62.4	6.1	1101	229	AL069706 Drosophila
6	61.8	6.0	1001	229	AL105023 Drosophila
7	61.8	6.0	1101	229	AL069706 Drosophila
8	61.2	5.9	1101	229	AL108721 Drosophila
9	60.2	5.9	1101	229	AL068307 Drosophila
10	59.6	5.8	1101	229	AL068307 Drosophila
11	59.2	5.8	554	137	BE600990 p11_95_F0
12	58.4	5.7	842	229	AL055007 Drosophila
13	58	5.6	781	229	AL053444 Drosophila
14	58	5.6	996	229	AL071063 Drosophila
15	56.8	5.5	1218	166	BE427889 PFR6899 I
16	56.6	5.5	1201	229	AL106652 Drosophila
17	56.2	5.5	1101	229	AL063932 Drosophila
18	56	5.4	945	231	AL285149 Tetradon

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

## ALIGNMENTS

### FEATURES

#### Location/Qualifiers

Query Match

Best Local Similarity 36.1%; Pred. No. 3-3e-06;  
Matches 208; Conservative 90; Mismatches 273; Indels 5; Gaps 2

	QY	450	TAAACAGCGAATGTAATGACCAAGTCGAATGATTACTAACAAAGAGTTAAACAGA	509
	Db <td>524</td> <td>TAAAAAACTCMMTATAAACACVAMCACACCCCMAAACTTCCAAAATAAAAAAAAAAW</td> <td>583</td>	524	TAAAAAACTCMMTATAAACACVAMCACACCCCMAAACTTCCAAAATAAAAAAAAAAW	583
	QY <td>510</td> <td>TGTCTATTAAGTTAAACACTGCACAGACGACCATGTTAACTGGAAAGCGTACAGTTGAAT</td> <td>569</td>	510	TGTCTATTAAGTTAAACACTGCACAGACGACCATGTTAACTGGAAAGCGTACAGTTGAAT	569
	Db <td>584</td> <td>YWCCTTAATTCCTCAACAACAAAATTWACATTTTMCCTCAAAAYCAAACCTCCATWCAWW</td> <td>643</td>	584	YWCCTTAATTCCTCAACAACAAAATTWACATTTTMCCTCAAAAYCAAACCTCCATWCAWW	643
	QY <td>570</td> <td>AGATGATCCACAAATATCTACTACAGTATTCAGATGGAAAAAACCCTATNAAGTAATATT</td> <td>629</td>	570	AGATGATCCACAAATATCTACTACAGTATTCAGATGGAAAAAACCCTATNAAGTAATATT	629
	Db <td>644</td> <td>TTTTYAAAAAAAYTYTCCAAAAAMTTMATTTTATATYMAIYCACCCHAAAAATTTTATT</td> <td>703</td>	644	TTTTYAAAAAAAYTYTCCAAAAAMTTMATTTTATATYMAIYCACCCHAAAAATTTTATT	703
	QY <td>630</td> <td>GTCITAGTCTGGGAATATTTTCATTTTAAATCAGCAAAATTTATCAAGATGAATCAACACCAAT</td> <td>689</td>	630	GTCITAGTCTGGGAATATTTTCATTTTAAATCAGCAAAATTTATCAAGATGAATCAACACCAAT	689
	Db <td>704</td> <td>TTTTWAAAYAAAAAAMMWMAAAAAAAATATTTTWTWMAAA--AAWYMAAACCTT</td> <td>761</td>	704	TTTTWAAAYAAAAAAMMWMAAAAAAAATATTTTWTWMAAA--AAWYMAAACCTT	761
	QY <td>690</td> <td>TTTGATATATNTCTGANAATCCAAATTTTAAACAAGCATCAACACATATTTGAATTTTAA</td> <td>749</td>	690	TTTGATATATNTCTGANAATCCAAATTTTAAACAAGCATCAACACATATTTGAATTTTAA	749
	Db <td>762</td> <td>TTYAATMTAAWMYCAAAYYYYCYAAAAAAAHAHTWMAAAAAHWTTTTTAAATAAAT</td> <td>821</td>	762	TTYAATMTAAWMYCAAAYYYYCYAAAAAAAHAHTWMAAAAAHWTTTTTAAATAAAT	821
	QY <td>750</td> <td>CCTGAGCTCTGTGATTTAAACACAACTCTCCACATTTATATTAAGAGAGGTTGGAAC</td> <td>809</td>	750	CCTGAGCTCTGTGATTTAAACACAACTCTCCACATTTATATTAAGAGAGGTTGGAAC	809
	Db <td>822</td> <td>TATTTTAATTAATAATTTAAAAAAMAAAAAMTTAAATCTTTCAAAATATCAAAAAT</td> <td>881</td>	822	TATTTTAATTAATAATTTAAAAAAMAAAAAMTTAAATCTTTCAAAATATCAAAAAT	881
	QY <td>810</td> <td>TTTGTCTAGTCGAGCAGGCTCCAACCCAGTACTCAGAAATTCCTCATCTATTAATAGATGA</td> <td>869</td>	810	TTTGTCTAGTCGAGCAGGCTCCAACCCAGTACTCAGAAATTCCTCATCTATTAATAGATGA	869
	Db <td>882</td> <td>WTTTWMAVCYTTLAAMAAYCACGYACCTAAATWMAVCHTTTMAATTTMAATWMMWMT</td> <td>941</td>	882	WTTTWMAVCYTTLAAMAAYCACGYACCTAAATWMAVCHTTTMAATTTMAATWMMWMT	941
	QY <td>870</td> <td>ATTATTTCTCTATTTATGCCCGCAAAATTAATTTGGCGATCTGGAATTAATCAATTTATCA</td> <td>929</td>	870	ATTATTTCTCTATTTATGCCCGCAAAATTAATTTGGCGATCTGGAATTAATCAATTTATCA	929
	Db <td>942</td> <td>TTTTWAATTTMAAYTTCYYCTTCANMAAAAAAMTAATTTATTAHYHMAAAATATYYWHT</td> <td>1001</td>	942	TTTTWAATTTMAAYTTCYYCTTCANMAAAAAAMTAATTTATTAHYHMAAAATATYYWHT	1001
	QY <td>930</td> <td>AACAATATGATGATTTGATGATCAGACGATGCCAACCATTTGAATTTGTCATTCGAGGT</td> <td>989</td>	930	AACAATATGATGATTTGATGATCAGACGATGCCAACCATTTGAATTTGTCATTCGAGGT	989
	Db <td>1002</td> <td>AAAAAAATMTMAA---AAWYTTAAMAAYCAAAAAATTTTCTTTTCCCAAAWW</td> <td>1058</td>	1002	AAAAAAATMTMAA---AAWYTTAAMAAYCAAAAAATTTTCTTTTCCCAAAWW	1058
	QY <td>990</td> <td>ATTAAATCAAAATGTTAAATTAACCTTAACGAAGA</td> <td>1025</td>	990	ATTAAATCAAAATGTTAAATTAACCTTAACGAAGA	1025
	Db <td>1059</td> <td>WHAAAAAHMAYCMWYTTTCMACTCTTTTMAAA</td> <td>1094</td>	1059	WHAAAAAHMAYCMWYTTTCMACTCTTTTMAAA	1094
RESULT	2			
LOCUS	AV426745			
DEFINITION	AV426745 lotus japonicus young plants (two-week old) Lotus japonicus cDNA clone MM01b02_r_5', mRNA sequence.			
ACCESSION	AV426745			
VERSION	AV426745.1			
SOURCE	EST.			
ORGANISM	Lotus japonicus. Lotus japonicus Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta Magnoliophyta; eudicotyledons, core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotus. 1 (bases 1 to 420) Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S. Generation of 7137 non-redundant expressed sequence tags from a legume, Lotus japonicus DNA Res. 7 (2), 127-130 (2000)			
JOURNAL	MEDLINE			
COMMENT	Contact: Yasukazu Nakamura The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yana 1533-3, Kisarazu, Chiba 292-0812, Japan Email: ynakamu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/ Location/Qualifiers			
FEATURES	Source			
	1..420			

```

/clone.lib="Lotus japonicus young plants (two-week old)"
/dev_stage="young plants (two-week old)"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI; isolate=Miyakojima M3-20"
BASE COUNT      98 a      72 c      132 g      118 t
ORIGIN

```

Query Match	6.6%	Score 68.2	DB 30	Length 420
Best Local Similarity	55.8%	Pred. No. 9.7e-06		
Matches 130	Conservative 0	Mismatches 103	Indels 0	Gaps 0

[illegible]

Accession	LOCUS	DEFINITION	LOCUS	DEFINITION
RESULTS 3	CNS00396/c	CNS00396	1101 bp	DNA
		Drosophila melanogaster genome survey sequence TERT3 end of BAC #		GSS
		BACR08K10 of RPc1-98 library from Drosophila melanogaster (fruit		fly), genomic survey sequence.

VERSION	AL0063921.1	GI:4941778
KEYWORDS	GSS.	
SOURCE	fruit fly.	
ORGANISM	Drosophila melanogaster	
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta	
	Pterygota; Neoptera; Endopterygota; Diptera; Brachyera;	
	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
REFERENCE	1 (bases 1 to 1101)	

**JOURNAL** Submitted (02-JUN-1999) Genoscope - Centre National de Séquençage  
BP 191 91006 Evry cedex - FRANCE (E-mail : segretef@genoscope.cns.fr)  
- Web : www.genoscope.cns.fr

**COMMENT** Determination of this BAC-end sequence was carried out as part of a

The BDGP is constructing a physical map of the *Drosophila melanogaster* genome using these BACs. For further information please see [http://www.fruitfly.org/TheBDGP/Drosophila\\_melanogaster\\_BAC\\_library](http://www.fruitfly.org/TheBDGP/Drosophila_melanogaster_BAC_library) was prepared by Kazuhiro Oosagawa and Aaron Mammosser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPT-98 and was constructed by partial EcoRI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

Location/Qualifiers

BASE COUNT	201 a	64 c	131 g	202 t	503 others
ORGANISM="Drosophila melanogaster"					
/db_xref="taxon:7227"					
/clone_lib="RPCT-98"					
/clone="BACROBK10"					
/note="end : TET3"					

Query Match	6.1%;	Score 63.2;	DB 229;	Length 1101;
Best Local Similarity	17.4%;	Pred. No. 0.00015;		
Matches 110;	Conservative 276;	Mismatches 244;	Indels 4;	Gaps 2;

OY	189	TATTACGTTAACAGCCATGACTCACTTATTTGGTTCACACACGACCCGCTGTTAACAAATAT	248
Db	1100	KARRGGDDTWTMRDTRKCDMDWTKKMTWTKWKDADDRRRAGADRRAMDDGAGTWTATATW	1040
OY	249	TGATTGTAGATAGCAAAAGTAGATATGCGACACAAAGACAAATTCGTTAGACACATGG	3080
Db	1040	WWWATTTDDTWTWMDWWWATTAAKTDFTAWTWTWTAMRADMGWRDGAARDBDADTADGA	9810
OY	309	TGATGAGACGTTACGGGGCTACGGTATTTAGGGTGAATG--CGTGTATGATGAACGGGCA	3660
Db	960	GRRDGRRRKRRKDKDKDDDDKKGGKKKKAKAAKAAKMTATKMWDDMDKMKWGAOKRKA	9210
OY	367	TCACATTTATACCAAGACTTTTTTAAAGCAAAAGCAAGCAACGACCCACAATTTACAGTG	4260
Db	920	DDDDGAGDKDDGKGDADDDTDGTDKDDDKDKMDMKKGTWGDATWMAATDMMWNG	8610
OY	427	AAAGTATTCGCAAGTTTATGATGTAACACAGCCAGTATATGACACAAAGTCAATGGATT	4860
Db	860	WADADMTWMDAAADDMWADDDWMDAMWKMDDMAMTGARTADRRDWMGRRAKRGARRRRD	8010
OY	487	ACTAACAAGAAGGTTTAAACAAGATGCTATATGATTAAGACATCGACGACGAGAGTTTA	5460
Db	800	RKRADDKRDADDDDDAATWTWTTTTTTTTROTDDPKMKTTDTWTRAADRTDRDDDDDRDR	7410
OY	547	ACT--GGAAGACGTACAGTTGAAATTAGATGATGATCCACAATATCTACACGTAATCAAGATG	6040
Db	740	AGTGRKKRRRTTKRRMRKRRDTRMWDADADADTARDDRRRRDDGDADGKKGTGRKRRRD	6810
OY	605	GAAAAAACCCATATAAAGTATATATGCTGAAGTCGGGAATATTCATTTTAATCAGCAAA	6640
Db	600	RATMDTRDAMWADAMWTTDTDDMDKRDORRRKRRRRRTTARAAWMMWTKAWDMA	6210
OY	665	TTTATCAAGATGATCAACACCAATTTGGATATATCTGAAATATCCAAATTTTAACAAGCA	7240
Db	620	KWDKTRADRRDRRAADDTWIDARADDDMAKARWRARRDRRAARARADRRWTTIKGTTTA	5610
OY	725	ATCAACAACATATTTGAATTTATTTACTTGAAGTCCTGATTTTAACAACAATTCCTTCACA	7840
Db	560	TWTWTWAAAAAAMWAMWATATATTTTWTWTTTTTTTTTTTTTTTTAAWMAWMTATW	5010
OY	785	ATTATATATAAAGAGAGCTTGGCACTTGGCTAGT	818
Db	500	AAWTTAAWAAAAAAMAAATTTTTTTTTTTTTTWTWMT	467

RESULT	4
LOCUS	BE920828
DEFINITION	BE920828 469 bp mRNA EST
ACCESSION	EST4425197 potato leaves and petioles Solanum tuberosum cDNA clone.
VERSION	csfBI0P19 5' sequence, mRNA sequence.
KEYWORDS	BE920828 BE920828.1 GI:10446904
SOURCE	EST.
ORGANISM	potato.
REFERENCE	Solanum tuberosum Euhayota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterid I; Solanales; Solanales; Solanales; Solanum. 1 (bases 1 to 469)
AUTHORS	van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J., Uiterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning, C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
TITLE	Generation of ESTs from potato leaves and petioles
JOURNAL	Unpublished (2000)
COMMENT	Contact: Cathy Ronning The Institute for Genomic Research For clone request: please contact Research Genetics, Libraries



RESULT	7
CNS00EVL	
LOCUS	CNS00EVL 1101 bp DNA GSS 04-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC: BACR2B823 of RpC1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION	AL069706
VERSION	AL069706.1 GI:4943849
KEYWORDS	GSS.
SOURCE	Fruit fly.
ORGANISM	Drosophila melanogaster
REFERENCE	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
AUTHORS	1 (bases 1 to 1101)
TITLE	Genoscope.
	Direct Submission

QY	371	AATATATCCAGACTTTTAAACCAACCAAGCAACGACCCACAATATACGTAAG	430
Db	492	AATATAAAAAAAAAAATTTTTTMMWMTATTTTMMWMTTATTAATTAATAAAAAAATAT	551
QY	431	TATCTGCAAGTTTATAGATGTAACCAACGCAATGATATGACCAAGTCATGATTA	490
Db	552	TTAAMMAAATATMAATTAAMAAATTTTMAAMMTATATTAATWTATAAATWTMTTAA	611
QY	491	ACCAAGAGGTTAAACCAAGATGCTATTAAGTACACATCGACACGACGAGTAACTG	550
Db	612	AAAATATTTTTTMMWMTAAATTTTTTAAATTAATTAATTAATTAATTAATTAATTA	671
QY	551	GAAACGCTACAGTGTATGATGATCCAACTTACTACACGATTAACAGATGAGAAA	610
Db	672	ATTAAATTAATTTTMAATTAAMAAAAAAAAAAAAAAAAAATMAAATATATTAATTA	731
QY	611	ACCCTATAAAGTAATATTTGCTTAAGCTCGGATATATTCATTTTAATCAGCAATTTATC	670
Db	732	WAATTAAMMAAMATMMWMTAATATATMTATATATATMTATTAAMMTTMMWMTATTA	791
QY	671	AAGATGATCAACACCAATTTGGATATATCTGAAATCCAAATTTTAACAGCAATCAAA	730
Db	792	WAATTAAMMAAMWMTAATTAATMAATMAATMAAAMWMTATMMWMTATTAAMWMTAA	851
QY	731	CACATATGAATATTTACTGTGAGCTGTGATTTAACAACAATCTCAACAATTAAT	790
Db	852	AAAATTAATTAATMAATWATTAAMAAAAAATATMAATTTMTTTTMMWMTATTAAMWAT	911
QY	791	ATAAAMAGAGAGTTGGAACCTTGGCTAGTCAGACAGTCCAACACACTTCCAGAAATTC	850
Db	912	AMAAAMMAAAAAAAAAAATTAAMAMMTTATTTATTTAAMTWTATMTATTTATTTW	971
QY	851	CCATCTATTAATAGATGAATTAATTCGTATTAATGCCCCGAAATTAATTTGGCGATCTG	910
Db	972	AMTWTATAMTWTATATWTATATWTMTAMTAMATATATTTATTAAMWTATTTTAA	1031
QY	911	GAAATTTCAATTTTATCAACAACAATATGTGATTT	945
Db	1032	TAAATATMTATMMWMTAMATATTAAMMAATTAATTT	1066



[illegible][illegible]









**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2001, 10:38:44 ; Search time 66.69 Seconds  
(without alignments)  
2694.050 Million cell updates/sec

Title: US-09-376-633-1

Perfect score: 1029

Sequence: 1 ATGGAATTATCGCATTCACACT.....TAACTTTACGAAGAAGTCA 1029

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 302621 segs, 87301344 residues

605242

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA:\*

1: /cgn2\_6/prodata/2/lna/5A\_COMB.seq:\*

2: /cgn2\_6/prodata/2/lna/5B\_COMB.seq:\*

3: /cgn2\_6/prodata/2/lna/5A\_COMB.seq:\*

4: /cgn2\_6/prodata/2/lna/5B\_COMB.seq:\*

5: /cgn2\_6/prodata/2/lna/PCTUS\_COMB.seq:\*

6: /cgn2\_6/prodata/2/lna/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1029	100.0	1029	3	US-08-978-456-1
2	1029	100.0	1029	2	US-08-978-456-1
3	1029	100.0	1029	1	US-08-978-456-1
4	1029	100.0	1029	4	US-08-978-456-1
5	1029	100.0	1029	5	US-08-978-456-1
6	1029	100.0	1029	6	US-08-978-456-1
7	1029	100.0	1029	7	US-08-978-456-1
8	1029	100.0	1029	8	US-08-978-456-1
9	1029	100.0	1029	9	US-08-978-456-1
10	1029	100.0	1029	10	US-08-978-456-1
11	1029	100.0	1029	11	US-08-978-456-1
12	1029	100.0	1029	12	US-08-978-456-1
13	1029	100.0	1029	13	US-08-978-456-1
14	1029	100.0	1029	14	US-08-978-456-1
15	1029	100.0	1029	15	US-08-978-456-1
16	1029	100.0	1029	16	US-08-978-456-1
17	1029	100.0	1029	17	US-08-978-456-1
18	1029	100.0	1029	18	US-08-978-456-1
19	1029	100.0	1029	19	US-08-978-456-1
20	1029	100.0	1029	20	US-08-978-456-1
21	1029	100.0	1029	21	US-08-978-456-1
22	1029	100.0	1029	22	US-08-978-456-1
23	1029	100.0	1029	23	US-08-978-456-1
24	1029	100.0	1029	24	US-08-978-456-1
25	1029	100.0	1029	25	US-08-978-456-1
26	1029	100.0	1029	26	US-08-978-456-1
27	1029	100.0	1029	27	US-08-978-456-1

28	38	3.7	10968	2	US-08-680-327-2	Sequence 2, Appl
29	37.8	3.7	2058	2	US-08-749-391-1	Sequence 1, Appl
30	37.8	3.7	2058	4	US-09-390-200-1	Sequence 1, Appl
31	37.2	3.6	1258	4	US-08-335-518-1	Sequence 1, Appl
32	37.2	3.6	1258	4	US-08-988-054-1	Sequence 1, Appl
33	37.2	3.6	1258	4	US-08-988-054-1	Sequence 1, Appl
34	37.2	3.6	1258	4	US-08-988-054-1	Sequence 1, Appl
35	37.2	3.6	1258	4	US-08-988-054-1	Sequence 1, Appl
36	37.2	3.6	1258	4	US-08-988-054-1	Sequence 1, Appl
37	37.2	3.6	1258	4	US-08-988-054-1	Sequence 1, Appl
38	37.2	3.6	1258	4	US-08-988-054-1	Sequence 1, Appl
39	37.2	3.6	1258	4	US-08-988-054-1	Sequence 1, Appl
40	37.2	3.6	1258	4	US-08-988-054-1	Sequence 1, Appl
41	37.2	3.6	1258	4	US-08-988-054-1	Sequence 1, Appl
42	37.2	3.6	1258	4	US-08-988-054-1	Sequence 1, Appl
43	37.2	3.6	1258	4	US-08-988-054-1	Sequence 1, Appl
44	37.2	3.6	1258	4	US-08-988-054-1	Sequence 1, Appl
45	37.2	3.6	1258	4	US-08-988-054-1	Sequence 1, Appl

#### ALIGNMENTS

```
RESULT 1
US-08-978-456-1
Sequence 1, Application US/08978456
Patent No. 6010881
GENERAL INFORMATION:
APPLICANT: Palmer, Leslie M.
TITLE OF INVENTION: No. 6010881el ribc
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESS: Dechart Price & Rhoads
CITY: Philadelphia
STATE: PA
COUNTRY: US
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,456
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dickinson, Todd O
REGISTRATION NUMBER: 28,354
REFERENCE/DOCKET NUMBER: P50444-9
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2252
TELEFAX: 215-994-2222
TELEX:
INFORMATION FOR SEQ. ID NO. 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1029 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: Linear
US-08-978-456-1
Query Match 100.0%; Score 1029; DB 3; Length 1029;
Best Local Similarity 100.0%; Pred. No. 3.7e+246;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Db 1 ATGCATTTATGCGATTCACTTCAATATATGTCACAAAGTCACAAAGCTTATATCCACC 60  
QY 61 GTTGGCGCTGTGTAGTATATAGAGTAGAGTGTGTTGTTGTCACACTTGAGAAAA 120  
Db 61 GTTGGCGCTGTGTAGTATATAGAGTAGAGTGTGTTGTTGTCACACTTGAGAAAA 120  
QY 121 GGTGCAAGCATGCGAGGTTCAAGCAGTGTATGCGCAACAAATCTGAAGTGGC 180  
Db 121 GGTGCAAGCATGCGAGGTTCAAGCAGTGTATGCGCAACAAATCTGAAGTGGC 180  
QY 181 ACGATTTATTTAGCTTAGAGCCATGATGTCATTTTGTTCACACACCCTGTGTAGC 240  
Db 181 ACGATTTATTTAGCTTAGAGCCATGATGTCATTTTGTTCACACACCCTGTGTAGC 240  
QY 241 AAAATTATTTGATTTAGATAGCAAAAGTAGTATGCGCAACAAAGACATTCGTTAGC 300  
Db 241 AAAATTATTTGATTTAGATAGCAAAAGTAGTATGCGCAACAAAGACATTCGTTAGC 300  
QY 301 ACACATGCTGATGACGCTTAGCGGCTCAGGCTATGAGCTGATGCTGATGATGA 360  
Db 301 ACACATGCTGATGACGCTTAGCGGCTCAGGCTATGAGCTGATGATGATGA 360  
QY 361 CGGCGATCACAATTTATACCAAGACTTTTAAAGCAAAAGCAAGCACTGCCACAAT 420  
Db 361 CGGCGATCACAATTTATACCAAGACTTTTAAAGCAAAAGCAAGCACTGCCACAAT 420  
QY 421 ACAGTGAAGATGTCGCAAGTGTAGTGTAAACAAGCAAGATATGAGCAAAAGTCAA 480  
Db 421 ACAGTGAAGATGTCGCAAGTGTAGTGTAAACAAGCAAGATATGAGCAAAAGTCAA 480  
QY 481 TGGATTACCAAGAGGTTAACAAGATGCTATAGCTTAAGTACATGACGACGCGA 540  
Db 481 TGGATTACTAACAAGAGGTTAACAAGATGCTATAGCTTAAGTACATGACGACGCGA 540  
QY 541 GTTGAATCTGGAAGAGCTACAGTGAATATGATGATCCAAATATACCTACAGTATCA 600  
Db 541 GTTGAATCTGGAAGAGCTACAGTGAATATGATGATCCAAATATACCTACAGTATCA 600  
QY 601 GATGCAAAAACCCCTATTAAGTATATGCTAAAGTCTGGCAATATTCATTTAATCAG 660  
Db 601 GATGCAAAAACCCCTATTAAGTATATGCTAAAGTCTGGCAATATTCATTTAATCAG 660  
QY 661 CAAATTTATCAAGATGAATCAACACCAATTTGATATATACGAAATCCAAATTAACA 720  
Db 661 CAAATTTATCAAGATGAATCAACACCAATTTGATATATACGAAATCCAAATTAACA 720  
QY 721 AGCAATCAACACATATGGAATTTATTTACTGGAAGTCTGTGATTTAACAACATTC 780  
Db 721 AGCAATCAACACATATGGAATTTATTTACTGGAAGTCTGTGATTTAACAACATTC 780  
QY 781 CACATTTATATTAAGAGAGGAGTGGAACTTGTCTAGTGGAGGAGGTCACACCTACT 840  
Db 781 CACATTTATATTAAGAGAGGAGTGGAACTTGTCTAGTGGAGGAGGTCACACCTACT 840  
QY 841 TCAGATTTCTCATCTATTTATATAGATGAATTTATCTATTTATGCCCCGAAATTAAT 900  
Db 841 TCAGATTTCTCATCTATTTATATAGATGAATTTATCTATTTATGCCCCGAAATTAAT 900  
QY 901 GCGGATCTGGAATTTATCAATTTTATCAACAAATGATGATGATGATGATGATGCG 960  
Db 901 GCGGATCTGGAATTTATCAATTTTATCAACAAATGATGATGATGATGATGATGCG 960  
QY 961 AACCAATTTGAATTTGTTATTCGAGTTATTAATCAAAAGTTAAATTAACCTTACGA 1020  
Db 961 AACCAATTTGAATTTGTTATTCGAGTTATTAATCAAAAGTTAAATTAACCTTACGA 1020  
QY 1021 AAGAGTGA 1029  
Db 1021 AAGAGTGA 1029

RESULT 2  
US-08-977-554-7

Sequence 7, Application US/08977554  
Patent No. 5891672  
GENERAL INFORMATION:  
APPLICANT: Palmer, Leslie M.  
APPLICANT: Pedon, Jason C.  
APPLICANT: Warren, Richard L.  
APPLICANT: Trainl, Christopher M.  
APPLICANT: Wang, Min  
APPLICANT: Jaworski, Deborah D.  
APPLICANT: Mooney, Jeffrey  
APPLICANT: Debouck, Christine  
APPLICANT: Zhong, Yixi  
APPLICANT: Black, Michael  
TITLE OF INVENTION: rDNA  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESS: Dechert, Price & Rhoads  
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103-2793  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/977,554  
FILING DATE:  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US97/02318  
FILING DATE: 19-FEB-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Dickinson, O. Todd  
REGISTRATION NUMBER: 28,354  
REFERENCE/DOCKET NUMBER: P50444-07  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215/994-2252  
TELEFAX: 215/994-2222  
TELEX:  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3336 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-977-554-7  
Query Match 100.0%; Score 1029; DB 2; Length 3336;  
Best Local Similarity 100.0%; Pred. No. 5.3e-246;  
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGCATTTATGCGATTCACTTCAATATATGTCACAAAGTCACAAAGCTTATATCCACC 60  
Db 1 ATGCATTTATGCGATTCACTTCAATATATGTCACAAAGTCACAAAGCTTATATCCACC 60  
QY 61 GTTGGCGCTGTGTAGTATATAGAGTAGAGTGTGTTGTTGTCACACTTGAGAAAA 120  
Db 61 GTTGGCGCTGTGTAGTATATAGAGTAGAGTGTGTTGTTGTCACACTTGAGAAAA 120  
QY 121 GGTGCAAGCATGCGAGGTTCAAGCAGTGTATGCGCAACAAATCTGAAGTGGC 180  
Db 121 GGTGCAAGCATGCGAGGTTCAAGCAGTGTATGCGCAACAAATCTGAAGTGGC 180  
QY 181 ACGATTTATTTAGCTTAGAGCCATGATGTCATTTTGTTCACACACCCTGTGTAGC 240  
Db 181 ACGATTTATTTAGCTTAGAGCCATGATGTCATTTTGTTCACACACCCTGTGTAGC 240  
QY 241 AAAATTATTTGATTTAGATAGCAAAAGTAGTATGCGCAACAAAGACATTCGTTAGC 300  
Db 241 AAAATTATTTGATTTAGATAGCAAAAGTAGTATGCGCAACAAAGACATTCGTTAGC 300



Db	241	AAATTTATGATCTGTAAGATAGCAAAAAGTAGTATACCAACAAAAAGCAATCTGTTAGAC	300
Qy	301	ACACATGCTGATGAGACGTTACGGGCTACCGGATTTGAGGTTGAATCGTTGATGATGA	360
Db	301	ACACATGCTGATGAGACGTTACGGGCTACCGGCTATGAGGTTGAATCGTTGATGATGA	360
Qy	361	CGGGCATACAAATTAATCCAGACTTTTAAAAACCAAAAGCAACACGCCACAATT	420
Db	361	CGGGCATACAAATTAATCCAGACTTTTAAAAACCAAAAGCAACACGCCACAATT	420
Qy	421	ACAGTGAAGATATGTCNAAGTTTATGATGTAACACAGCATGATTAATGACAAATGCA	480
Db	421	ACAGTGAAGATATGTCNAAGTTTATGATGTAACACAGCATGATTAATGACAAATGCA	480
Qy	481	TGATTTACTAACAAGAGGTTAAACAAGATGCTATTAAGTTAAGACATGACACGCGCA	540
Db	481	TGATTTACTAACAAGAGGTTAAACAAGATGCTATTAAGTTAAGACATGACACGCGCA	540
Qy	541	GTGTTAACTGGAACGATACGTTGATTAATGATGATCCACAATATCTACAGCATTTCCA	600
Db	541	GTGTTAACTGGAACGATACGTTGATTAATGATGATCCACAATATCTACAGCATTTCCA	600
Qy	601	GATGGAAAAAACCCTATATAAAGTATATTTGCTAAGCTGGGAATATTCATTTTAATCAG	660
Db	601	GATGGAAAAAACCCTATATAAAGTATATTTGCTAAGCTGGGAATATTCATTTTAATCAG	660
Qy	661	CAAAATTTATCAAGTGAATCAACACCAATTTGGATATATACGTGAATAATCCAAATTTAACA	720
Db	661	CAAAATTTATCAAGTGAATCAACACCAATTTGGATATATACGTGAATAATCCAAATTTAACA	720
Qy	721	AGCATCAACACATATTTGAATTAATTTACTGTGAAGCTTGATGATTTAACACAATTCCTT	780
Db	721	AGCATCAACACATATTTGAATTAATTTACTGTGAAGCTTGATGATTTAACACAATTCCTT	780
Qy	781	CACAAATTTATATAAAGAGAGAGAGTTGGAACCTTGCTAGTCGAGGAGGTCACACAGTACG	840
Db	781	CACAAATTTATATAAAGAGAGAGAGTTGGAACCTTGCTAGTCGAGGAGGTCACACAGTACG	840
Qy	841	TCAGAAATTCCTCATCTATATATATGATGAATTAATTTCTCTATTAATGCCCGAAATTAAT	900
Db	841	TCAGAAATTCCTCATCTATATATATGATGAATTAATTTCTCTATTAATGCCCGAAATTAAT	900
Qy	901	GCGCGATCTGCAAAATTTATCAATTTTATCAACAACAAATATATGATGATGACAGATGCG	960
Db	901	GCGCGATCTGCAAAATTTATCAATTTTATCAACAACAAATATATGATGATGACAGATGCG	960
Qy	961	AACCAATTTGAAATTTGTTCAATTCGAGATTAATTAATCAAAATGTTAAATTAATTTACGA	1020
Db	961	AACCAATTTGAAATTTGTTCAATTCGAGATTAATTAATCAAAATGTTAAATTAATTTACGA	1020
Qy	1021	AAGAAGTCA 1029	
Db	1021	AAGAAGTCA 1029	

RESULT 3  
 US-088-978-456-7  
 : Sequence 7, Application US/08978456  
 : Patent No. 6010881  
 : GENERAL INFORMATION:  
 : APPLICANT: Palmer, Leslie M.  
 : TITLE OF INVENTION: No. 6010881a1 r1bG  
 : NUMBER OF SEQUENCES: 7  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Dechert Price & Rhoads  
 : STREET: 4000 Bell Atlantic Tower, 1717 Arch Street  
 : CITY: Philadelphia  
 : STATE: PA  
 : COUNTRY: US  
 : ZIP: 19103  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Diskette  
 : COMPUTER: IBM Compatible

?	OPERATING SYSTEM:	DOS
?	SOFTWARE:	Fasteq for Windows Version 2.0
?	CURRENT APPLICATION DATA:	
?	APPLICATION NUMBER:	US/08/978.456
?	FILING DATE:	
?	CLASSIFICATION:	435
?	PRIOR APPLICATION DATA:	
?	APPLICATION NUMBER:	
?	FILING DATE:	
?	ATTORNEY/AGENT INFORMATION:	
?	NAME:	Dickinson, Todd O
?	REGISTRATION NUMBER:	28, 354
?	REFERENCE/DOCKET NUMBER:	P50444-9
?	TELECOMMUNICATION INFORMATION:	
?	TELEPHONE:	215-994-2252
?	TELEFAX:	215-994-2222
?	TELEX:	
?	INFORMATION FOR SEQ ID NO:	7:
?	SEQUENCE CHARACTERISTICS:	
?	LENGTH:	3336 base pairs
?	TYPE:	nucleic acid
?	STRANDEDNESS:	double
?	TOPOLOGY:	Linear
?	US-08-978-456-7	
Query Match	100.0%; Score 1029; DB 3; Length 3336;	
Best Local Similarity	100.0%; Pred. No. 5.3e-246;	
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0		
OY	1 ATGCAATTCGCATTCACCTTCCAAATATGTACAGAGTCAAAGCGTGTAATCCACC	60
DB	1 ATGCAATTCGCATTCACCTTCCAAATATGTACAGAGTCAAAGCGTGTAATCCACC	60
OY	61 GTGGCGCTGTGTGTTAATGAAGGTAGGATTTGCTATGTGGCCACTTGGAAAA	120
DB	61 GTGGCGCTGTGTGTTAATGAAGGTAGGATTTGCTATGTGGCCACTTGGAAAA	120
OY	121 GGTCACAAGCATGCGAGAGTTCAACACACTTGATATGGCACACAAAATCTGNAAGTGG	180
DB	121 GGTCACAAGCATGCGAGAGTTCAACACACTTGATATGGCACACAAAATCTGNAAGTGG	180
OY	181 ACGATTTATATTACGTTAGAGCCATGTAGTCATTTGGTTCAACACACCCTGTGTTAAC	240
DB	181 ACGATTTATATTACGTTAGAGCCATGTAGTCATTTGGTTCAACACACCCTGTGTTAAC	240
OY	241 AAAATTATTGATTTGAAGTAGCAAAAGTAGTATACGCACAAAAGACATTCGTTAGAC	300
DB	241 AAAATTATTGATTTGAAGTAGCAAAAGTAGTATACGCACAAAAGACATTCGTTAGAC	300
OY	301 ACACATGGTGAATGAGCGTTACGGGCTCAGCGTATTGAGCGTGAATGGCTGATGTGAA	360
DB	301 ACACATGGTGAATGAGCGTTACGGGCTCAGCGTATTGAGCGTGAATGGCTGATGTGAA	360
OY	361 CGGCGCATCACAATTTATCCAAAGACTTTTTTAAAGCAAAAGCAAGCAATGCCACAAATT	420
DB	361 CGGCGCATCACAATTTATCCAAAGACTTTTTTAAAGCAAAAGCAAGCAATGCCACAAATT	420
OY	421 ACAAGTGAAGTAGTCTGCAAGTTTAGATGTATAACAAGAGCATGATATGACAAAGTCAA	480
DB	421 ACAAGTGAAGTAGTCTGCAAGTTTAGATGTATAACAAGAGCATGATATGACAAAGTCAA	480
OY	481 TGGATTACTTAACAAGAGGTTAAACAAGAATCTAATAGTTAAGCATGACACAGCGCA	540
DB	481 TGGATTACTTAACAAGAGGTTAAACAAGAATCTAATAGTTAAGCATGACACAGCGCA	540
OY	541 GTGTTAACTGGAAGACGTACAGTTGAATAGATGATCCCAATATATCTACAGTATTTCAA	600
DB	541 GTGTTAACTGGAAGACGTACAGTTGAATAGATGATCCCAATATATCTACAGTATTTCAA	600
OY	601 GATGGAAGAAACCTATATAAGTAATATTGTCTAAGTCTGGGAATATTCATTTAATCAG	660
DB	601 GATGGAAGAAACCTATATAAGTAATATTGTCTAAGTCTGGGAATATTCATTTAATCAG	660

QY 661 CAAATTTATCAAGATGATCAACACCAATTTGGATATATCTGAAATCCAAATTTACA 720  
| | | | |  
DB 661 CAAATTTATCAAGATGATCAACACCAATTTGGATATATCTGAAATCCAAATTTACA 720  
QY 721 AGCAATCAACACATATTTGAATTTATTTCTGAGTCTGTGATTACCAATCTT 780  
| | | | |  
DB 721 AGCAATCAACACATATTTGAATTTATTTCTGAGTCTGTGATTACCAATCTT 780  
QY 781 CACAATTTATATATAAGAGAGAGTGGAACTTGTCTAGTCGAGGACAGCTCAACACTACT 840  
| | | | |  
DB 781 CACAATTTATATATAAGAGAGTGGAACTTGTCTAGTCGAGGACAGCTCAACACTACT 840  
QY 841 TCAGAAATTCCTCATCTATATATATAGATGATTTCTCTATATATGCCCCGAATTAAT 900  
| | | | |  
DB 841 TCAGAAATTCCTCATCTATATATATAGATGATTTCTCTATATATGCCCCGAATTAAT 900  
QY 901 GCGGATCTGGAATTTATCAATTTATCAACAAATGATGATGATGATGATGATGATG 960  
| | | | |  
DB 901 GCGGATCTGGAATTTATCAATTTATCAACAAATGATGATGATGATGATGATGATG 960  
QY 961 AACCAATTTGAATTTGTCATTCGAGTATTAATCAAAATGTTAAATTAATTAATTA 1020  
| | | | |  
DB 961 AACCAATTTGAATTTGTCATTCGAGTATTAATCAAAATGTTAAATTAATTAATTA 1020  
QY 1021 AAGAAGTGA 1029  
| | | | |  
DB 1021 AAGAAGTGA 1029

RESULT 4  
US-09-225-967-7  
Sequence 7, Application US/09225967  
Patent No. 6171598  
GENERAL INFORMATION:  
APPLICANT: Palmer, Leslie M.  
APPLICANT: Fedon, Jason C.  
APPLICANT: Warren, Richard L.  
APPLICANT: Traini, Christopher M.  
APPLICANT: Wang, Min  
APPLICANT: Jaworski, Deborah D.  
APPLICANT: Mooney, Jeffrey  
APPLICANT: Debourck, Christine  
APPLICANT: Zhong, Yixi  
APPLICANT: Black, Michael  
TITLE OF INVENTION: r1bA  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dechert, Price & Rhoads  
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103-2793  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/225.967  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/977.554  
FILING DATE:  
APPLICATION NUMBER: PCT/US97/02318  
FILING DATE: 19-FEB-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Dickinson, O. Todd  
REGISTRATION NUMBER: 28,354  
REFERENCE/DOCKET NUMBER: P50444-07  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 215/994-2252  
TELEFAX: 215/994-2222  
TELEX:  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3336 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-09-225-967-7

Query Match 100.0%; Score 1029; DB 4; Length 3336;  
Best Local Similarity 100.0%; Pred. No. 5.3e-246;  
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAATTAATGCGATTCACCAATTCGAATATGATGACAAAGTCAACAGTCTTAATCACACC 60  
| | | | |  
DB 1 ATGGAATTAATGCGATTCACCAATTCGAATATGATGACAAAGTCAACAGTCTTAATCACACC 60  
QY 61 GTTGGCGCTGTTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 120  
| | | | |  
DB 61 GTTGGCGCTGTTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 120  
QY 121 GGTGACAGCATGCGGAGGTTCAAGCATTGATATGCGCAACAAATGCTGAAGTGG 180  
| | | | |  
DB 121 GGTGACAGCATGCGGAGGTTCAAGCATTGATATGCGCAACAAATGCTGAAGTGG 180  
QY 181 ACGATTTATATATACGTTAGAGCAGTACTGATGATGATGATGATGATGATGATGATG 240  
| | | | |  
DB 181 ACGATTTATATATACGTTAGAGCAGTACTGATGATGATGATGATGATGATGATGATG 240  
QY 241 AAAATTTATGATTTGTAAGTATGACAAAGTATACGCAACAAAGACAAATTCGTTAGAC 300  
| | | | |  
DB 241 AAAATTTATGATTTGTAAGTATGACAAAGTATACGCAACAAAGACAAATTCGTTAGAC 300  
QY 301 AACATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
| | | | |  
DB 301 AACATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
QY 361 CCGGCAATCAATTTATACCAAGCTTTTAAAGCAAAAGCAACCACTGCCCAAAAT 420  
| | | | |  
DB 361 CCGGCAATCAATTTATACCAAGCTTTTAAAGCAAAAGCAACCACTGCCCAAAAT 420  
QY 421 ACGTGAAGATATCTGCAAGTATGATGATGATGATGATGATGATGATGATGATGATG 480  
| | | | |  
DB 421 ACGTGAAGATATCTGCAAGTATGATGATGATGATGATGATGATGATGATGATGATG 480  
QY 481 TGGATTACTAACAAGAGGTTTAAACAAGTCTTAAGTTAAGATGATGATGATGATGATG 540  
| | | | |  
DB 481 TGGATTACTAACAAGAGGTTTAAACAAGTCTTAAGTTAAGATGATGATGATGATGATG 540  
QY 541 GTGTTAACTGAAGAGCTACAGTTGAATATGATGATGATGATGATGATGATGATGATG 600  
| | | | |  
DB 541 GTGTTAACTGAAGAGCTACAGTTGAATATGATGATGATGATGATGATGATGATGATG 600  
QY 601 GATGAAAAAACCTATATAAGTATATGCTAGCTGCGGAATATTCATTTAATCAG 660  
| | | | |  
DB 601 GATGAAAAAACCTATATAAGTATATGCTAGCTGCGGAATATTCATTTAATCAG 660  
QY 661 CAAATTTATCAAGATGATCAACACCAATTTGGATATATATGAGAAATCCAAATTTACA 720  
| | | | |  
DB 661 CAAATTTATCAAGATGATCAACACCAATTTGGATATATATGAGAAATCCAAATTTACA 720  
QY 721 AGCAATCAACACATATTTGAATTTATTTCTGAGTCTGTGATTACCAATCTT 780  
| | | | |  
DB 721 AGCAATCAACACATATTTGAATTTATTTCTGAGTCTGTGATTACCAATCTT 780  
QY 781 CACAATTTATATATAAGAGAGTGGAACTTGTCTAGTCGAGGACAGCTCAACACTACT 840  
| | | | |  
DB 781 CACAATTTATATATAAGAGAGTGGAACTTGTCTAGTCGAGGACAGCTCAACACTACT 840  
QY 841 TCAGAAATTCCTCATCTATATATATAGATGATTTATCTCTATATATGCCCCGAATTAAT 900  
| | | | |



TELEX:  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3417 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-978-458-7

Query Match 10.9%; Score 112.6; DB 2; Length 3417;  
Best Local Similarity 49.2%; Pred No. 4,1e-19;  
Matches 386; Conservative 0; Mismatches 389; Indels 10; Gaps 3;

```
OY 1 ATGGAATTATGCGATTCACACTGCAATATGTACAAAGTCAAGGTTTATCCACC 60
DB 19 ATGAATTATGCAATAAACTGGCACAAAAGGGCGTGTACGTCAATCCCAATCTT 78
OY 61 GTTGGCGCTGTTGTAGTAAATGAGTATGTTGTTAGTGTGACACACTTGAGAAA 120
DB 79 GTTGGCGCAATTATGTAAAGATATACATTATCGCAAGGTTATCATGATTTTTT 138
OY 121 GGTGACAAAGCATGCGAGGTTTCAAGCACTGTATATGCGACAACAAAATGCTGAAGTGGC 180
DB 139 GGTGGCCCAATGCTGAGAGAAATGCTTTAAACTGTGAAATCCCTGTGAGCG 198
OY 181 ACGATTATATTTACGTTAGACCGCATGTACTATTTTGTTCACACACCCCTGTGTTAC 240
DB 199 ACGCTTTATGTAACACTTGAACCCCTGTGTCTACTTCGGAAAAACACCTCCTGTATAGAT 258
OY 241 AAAATTATGATTTAGTAAGTACCAAAAGTATACGCAAAAGA-----CAAT 292
DB 259 GCTATATATGATAGTGTATTTACAGAGTAGTCAATTGGAAGCCTGACGTATATCTATT 318
OY 293 CCTTAGACACACATGTCATGAGAGTTCACGGCTCACCGTATTTGAGTTGAGTGGCTTG 352
DB 319 GTATCTGGAAGAAAGAGTAAGTACTGTGAGAAATTAATCTTCAAGTTACTGTGGAAT 378
OY 353 ATGATGACGCGCATCACAATTATACCAAGACTTTTAAAGCA-AAAGCAAGCAACTG 411
DB 379 TTAGAAATGAGTGTCTTAACTTAATAAAGTTTAAAGTATATACCGACATGTA 438
OY 412 CCACAATTAAGTGAAGTATCTGCAATTTAGATGTAAGCAAGCGAATATATGA 471
DB 439 CCTATGTTTATGAAATATGCAATGTCATATGAGATGAAATATCCCTAAACAAAT 498
OY 472 CAAAGTCATGATTTACTTAAACAAGAGTTTAAACAAGATGCTATTAAGTAAAGACATCGA 531
DB 499 CAATGCAATGATTTACTGAAGAGAAAGCAAGAAAGCATGTCATCAGTTAGACACAT 558
OY 532 CACGACGCACTTTAACTGGAAGAGCTACAGTTGAATTAAGATTCACAAATATCTACA 591
DB 559 GTTATGCAATATGATGAGTGTCAATCTGTATTCAGAGATTTTGTGTCAGATGT 618
OY 592 CGTATTCAGATGAAAGAAACCTTATAAAGTATATTTGTCTAAGTCGGAAATATCAT 651
DB 619 ACATTGGAAGAGGAAATATCTTATCCGTATCATATGCGATACACATTTACGAATCTCT 678
OY 652 TTTA-ATCAGCAAAATTTATCAAGATGAATCAACACCAATTTGATATATCTGAAATTC 710
DB 679 CTACTCTTAAATGTTAAACAGCAATGATATTTAAACTTACATTTCCACTTCTCT 738
OY 711 AAATTTAACAGCAATCAACACATATTTGAAATTTATTTACTGGAAGCTGTGATTTAAC 770
DB 739 GAAGCAAAATTAATGAAGTATATCAAAATCATGCTGTGAATATCTTCCATTAAG 798
OY 771 AACAA 775
DB 799 AAAAA 803
```

RESULT 7  
US-08-978-454-7

```
; Sequence 7, Application US/08978454
; Patent No. 6017728
; GENERAL INFORMATION:
; APPLICANT: Palmer, Leslie M.
; APPLICANT: Fedon, Jason C.
; APPLICANT: Warren, Richard L.
; APPLICANT: Kosmatka, Anna L.
; APPLICANT: Shilling, Lisa K.
; APPLICANT: Stodola, Robert K.
; APPLICANT: Knowles, David J. C.
; APPLICANT: Black, Michael T.
; APPLICANT: Hodgson, John E.
; APPLICANT: Nicholas, Richard O.
; TITLE OF INVENTION: ribh
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,454
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/911,503
; FILING DATE: 15-AUG-1997
; APPLICATION NUMBER: PCT/US97/14436
; FILING DATE: 15-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, O. Todd
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: P50533-11
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215/994-2252
; TELEFAX: 215/994-2222
;
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3417 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; US-08-978-454-7
```

Query Match 10.9%; Score 112.6; DB 3; Length 3417;  
Best Local Similarity 49.2%; Pred No. 4,1e-19;  
Matches 386; Conservative 0; Mismatches 389; Indels 10; Gaps 3;

```
OY 1 ATGGAATTATGCGATTCACACTGCAATATGTACAAAGTCAAGGTTTATCCACC 60
DB 19 ATGAATTATGCAATAAACTGGCACAAAAGGGCGTGTACGTCAATCCCAATCTT 78
OY 61 GTTGGCGCTGTTGTAGTAAATGAGTATGTTGTTAGTGTGACACACTTGAGAAA 120
DB 79 GTTGGCGCAATTATGTAAAGATATACATTATCGGACAAAGTATCATGATTTTTT 138
OY 121 GGTGACAAAGCATGCGAGGTTTCAAGCACTGTATATGCGACAACAAAATGCTGAAGTGGC 180
DB 139 GGTGGCCCAATGCTGAGAGAAATGCTTTAAAACTGTGAAATCCCTGTGAGCG 198
OY 181 ACGATTATATTTACGTTAGAGCATGTACTATTTTGTTCACACACCCCTGTGTTAC 240
DB 199 ACGCTTTATGTAACACTTGAACCCCTGTGTCTACTTCGGAAAAACACCTCCTGTATAGAT 258
```

```

QY 241 AAATTATGATTTAGATGACAAAGTACTATACGCAACAAAAG-----CAATT 292
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 259 GCTATTAATCGATGATGATTACAGAGATGTCATTCGAACTGTAATCTTAT 318
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 293 CGTTAGACACATGATGATGAGACGTTACGGGCTCAGCGTATGAGTGAATCCGTTG 352
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 319 GTATCTGAAAAGAGATTAAGATCTTGAGAAAATATCTTCAAGTTACTGTTGAATT 378
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 353 ATGATGACGGGACATCAATTAACCAAGACTTTTAAAGCA-AAAGCAAGCACTG 411
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 379 TTAGAAAATGATGCTTAATTAATAAAGTTTAAAGTATATTAACCAAGCATGTA 438
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 412 CCACAAATTAAGTGAAGTATCTGCAAGTTAGATGCTAAACAAGCAATGATATGA 471
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 439 CCCATGCTTTTATGAATATATGCAATGTCATGCAATGCAAAAATGACCACTAAACAAAT 498
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 472 CAAAGTCAATGATTAATCAACAAAGAGGTTAAACAGATGCTTAATTAAGCAATGCA 531
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 499 CAATCCAAATGATTAATCAAGAAAGAAAGCAAGAAAGCAATGTCATCACTTACGACACTAT 558
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 532 CACGACGAGTGTAACTGGAAGACGTACACGTTGAATGATGATCCAAATATATCA 591
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 559 GTTAGTCAATTAATGCTGAGGAGTCAATGCTTATTAAGACGATCCCTTGCTGACATGT 618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 592 CGTATTCAGATGAAAAAACCTTAATAAGTATATTTGCTAAGTCTGGAAATATTCAT 651
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 619 AGATTGGAGGAAGGAAAAATCTATCCGTATCATATGCGATACCAATTAAGCAATCTCT 678
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 652 TTTA-ATCAGCAATTTATCAAGATGAATCAACCACTTTGATATATCAAGAAATCC 710
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 679 CTTCACCTTAATATCGTAATAACAGCAATATGATTTAAACTTACATTCGCACTTCTCT 738
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 711 AAATTTAACAAGCAACAACATATTTGAATTTACTTGAAGCTTGTGATTTAAC 770
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 739 GAAGCAAAAATAAATGAAGCTATATCAATATGATGCTGGAATCTTCCATTAAG 798
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 771 AACAA 775
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 799 AAAAA 803
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 8
US-09-385-288-7
; Sequence 7, Application US/09385288
; Patent No. 6171835
; GENERAL INFORMATION:
; APPLICANT: Palmer, Leslie M.
; APPLICANT: Fedon, Jason C. L.
; APPLICANT: Warren, Richard L.
; APPLICANT: Kosmatka, Anna L.
; APPLICANT: Shilling, Lisa K.
; APPLICANT: Stodola, Robert K.
; APPLICANT: Knowles, David J. C.
; APPLICANT: Black, Michael T.
; APPLICANT: Hogson, John E.
; APPLICANT: Nicholas, Richard O.
; TITLE OF INVENTION: ribh
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/385,288
; FILING DATE:

```

```

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,454
; FILING DATE:
; APPLICATION NUMBER: 08/911,503
; FILING DATE: 15-AUG-1997
; APPLICATION NUMBER: PCT/US97/14436
; FILING DATE: 15-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, O. Todd
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: P50533-11
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215/994-2252
; TELEFAX: 215/994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3417 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-09-385-288-7

```

```

Query Match 10.9%; Score 112.6; DB 4; Length 3417;
Best Local Similarity 49.2%; Pred. No. 4,1e-19;
Matches 386; Conservative 0; Mismatches 389; Indels 10; Gaps 3;

```

```

QY 1 ATGATTTATGCGATTCACATTCGCAATATGCTACAGGTCAACAGGTTTAATCCACC 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 19 ATGAATATGCAATAAACCTGGCACAAAAGGGCTGGTTACGTCATCCCAATCCATATG 78
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 GTGGCGCTGTTAGTAAATGAAGGTGAGATGTTGTTGTTATGTTGTCACACTTGAGAAA 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 79 GTGGCGCAATTTATTTAAAGATATATCAATTTATCGACAAAGTTATATGATGATTTT 138
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 GGTGACAAGCATGGGAGGTTCAAGCACTGTATATGCAACAACAAATGCTGAAGTCG 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 139 GGTGGCCACATGCTGAGAGAAATGCTCTTAAACCTGTAAGAAATCCCTGTGGAGCG 198
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 ACGATTTATATGCTTATGAGCCATGTATGCTATTTGTTTCAACACCCCTGTGTAC 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 199 ACGCTTATGTACCACTTGAACCCCTGTGCTCACTCGGAAACCCCTCGTATATGAT 258
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 241 AAATTTATGATTTGAAGTACCAAAAGTATGATACGCAACAAAAG-----CAATT 292
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 259 GCTATATGATATGATGATTAACAGATGATGATGAGACCTTGAAGCTGTAATCTTAT 318
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 293 CGTTAGACACATGATGATGAGACGTTACGGGCTCAGCGTATGAGGTTGAATGCGTTG 352
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 319 GTATCTGAAAAGAGATTAAGACTCTTGAGAAATATCTTCAAGTTACTGTTGAATT 378
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 353 ATGATGAACGGGACATCAATTAATTAACAGCTTTTAAAGCA-AAAGCAAGCACTG 411
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 379 TTAGAAAATGAGTCTTAATTAATAAAGTTTAAAGTATTAATTAACCAAGCATGTA 438
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 412 CCACAAATTAAGTGAAGTATCTGCAAGTTAGATGCTAAACAAGCAATGATATGA 471
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 439 CCCATGCTTTTATGAATATGCAATGTCATGATGAGAAAATTAAGCACTTAACAAAT 498
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 472 CAAAGTCAATGATTAATCAACAAAGAGGTTAAACAGATGCTTAATTAAGCAATGCA 531
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 499 CAATCCAAATGATTAATCAAGAAAGAAAGCAAGAAAGCAATGTCATCACTTACGACACTAT 558
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 532 CACGACGAGTGTAACTGGAAGACGTACACTTGAATTAAGTATGATCCAAATATATCA 591
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 559 GTTAGTCAATTAATGCTGAGGAGTCAATGCTTATTAAGACGATCCCTTGCTGACATGT 618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 592 CGTATTCAGATGAAAAAACCTATAAAGATATGTTAAGTCTGGGAATATTCAT 651
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 619 AGATTGGAGGAAGGAAATCTTATCCGTATCATATGCGATACCAATTTACGAATCTCT 678
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Qy	652	TTTA-ATCAGCAAAATTTATCAAGATGATCAACGCCAAATTTGGTATATCTCAAAATCC	710
Db	679	CTTACCTCTAAATTCGTAAAAACGCAAAATGATATATAAATCTTACATTTGCCACTTCTCT	728
Qy	711	AAATTTACAGCAATCAACACATATTGAAATTTACTGTGAAGCTGTGATTTAAC	770
Db	739	GAAAGCAAAAATTAATAATGAAGCTATATCAAAATCATGTGGCTGTGAATATCTTCCATTAAG	798
Qy	771	AACAA	775
Db	799	AAAAA	803

RESULT 9  
US-08-741-327E-14  
; Sequence 14, Application US/08741327E  
; Patent No. 502525A

```

1  GENERAL INFORMATION:
2  APPLICANT: Fuller, et al.
3  TITLE OF INVENTION: Riboflavin mutants as vaccines
4  TITLE OF INVENTION: against Actinobacillus pleuropneumoniae
5  NUMBER OF SEQUENCES: 17
6  CORRESPONDENCE ADDRESS:
7  ADDRESSEE: G. Kenneth Smith
8  STREET: 300 S. Wacker Drive
9  CITY: Chicago
10 STATE: IL
11 COUNTRY: USA
12 ZIP: 60606
13
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: Floppy disk
16 COMPUTER: IBM PC compatible
17 OPERATING SYSTEM: PC-DOS/MS-DOS
18 SOFTWARE: ASCII
19
20 CURRENT APPLICATION DATA:
21 APPLICATION NUMBER: US/08/741,327E
22 FILING DATE: October 28, 1996
23 CLASSIFICATION: 424
24
25 ATTORNEY/AGENT INFORMATION:
26 NAME: G. Kenneth Smith
27 REGISTRATION NUMBER: 43,135
28 REFERENCE/DOCKET NUMBER: 97704-E
29
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: 312-913-0001
32 TELEFAX: 312-913-9808
33
34 INFORMATION FOR SEQ ID NO: 14:
35 SEQUENCE CHARACTERISTICS:
36 LENGTH: 1230 bases
37 TYPE: nucleic acid
38 STRANDEDNESS: single
39 TOPOLOGY: linear
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
97
```

Query Match	7.48;	Score	76.2;	DB	2;	Length	1230;
Best Local Similarity	47.68;	Pred. No.	3.1e-10;				
Matches	293;	Conservative	0;	Mismatches	313;	Indels	9;
						Gaps	2

QY	9	TGATATTCACACTTGGCAATATATGTTACACAGGTCAACAGCGTTTATATCCACCCTGTGGCCG	68
	165	TGCCATTGCACTGGCAAAACAGGTTTATAGGTGGACCAATCCCATCCGCTTGGCGTTG	224
Db	69	TGTTGTATGTTAATGAAGGTAGGATGTTGTGATATGTTGGACACACTTGAGAAAGGATGCACAA	128
QY	225	TGTATTTGTCAAAAACGTTGAAATCGTTGGCCGAAGTTTACCATGTAAAGAATTTGGTGATG	284
Db	129	GCATGCGGAGGTTCAAGCACTTGATATATGACACAACAAATGCTGAAGGTGCGACGATTTA	188
QY	285	GCATGCGGAGCACTAATCCGCTTTTACATTTGTAAGGAGAGATCTTCCGCGGCGACCTGCTTA	344
Db	189	TATATCGTTAGACCACTATGATGCTATTTTGGTTCAACACACACCCTGTGTATTAACAAATAT	248
QY	345	TGTATACCTGTAGGCTTGTGTGTATCAACGCGCCGACCCGCTGTGTTCGGAATTTATTTAT	404

QY	249	TGATGTGAAGTTACCAAAAGTAGTATACGCAACAAAGA---CAATTCGTTAGACACACA	305
Db	405	TGAAGAGGCATTAAAAAGATTATTACGGTTCGAGCGATCCGAAATCCTTTAGTACAGC	464
QY	306	TGGTG-----ATGAGACGTTACGGGCTACCGGTATTGAGGTGGAAATGCGTTGATGTA	359
Db	465	GCGGGGAGCAAAATCAGCTACGCCACAGCGCGCTGGAAGTGGTGAAGGTTACTCAAGA	524
QY	360	ACGGGCATCAACAATTATATCCAGAAGCTTTTAAAGCAAAAGCAAGCAACTGCCACAAAT	419
Db	525	AGAAATGTGATCGCTTAAACCCGATTTTTTTTCCACTATATTCCAATAAACGTCCGATGT	584
QY	420	TACAGTGAAGTATCTCGCAAGTTAGATGTGTAACCAAGCAAGATGATATGACAAAGTCA	479
Db	585	GCTATATGAATATATCCATGACGGCAGACGGCAAAATTGCAACCGGTATGCGGCAATCCAA	644
QY	480	ATGATTTACTAACAAAGAGGTTAAACAAGATGCTCTATATAGTTAAGACATCGACAGCAGC	539
Db	645	ATGATTTACCGGTGAATATCCAGAACAGCAAGTGTCACCAACCAACGTCATCAATATATGTGC	704
QY	540	AGCTGTTAATCTGGAAGACGTACAGTTGTAATTAGATGATCCAAATATATCTACACGTATTCA	599
Db	705	GATTATGTCGGGTAGATATGCTACTTCCGCATTAACCCGATGTTAAATATGCGCAATGCC	764
QY	600	AGATGGAACAAAAACC	614
Db	765	CAATGCCAAACACACC	779

RESULT 10  
US-08-232-463-14/c  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION.

APPLICANT: DORNER, F.  
APPLICANT: SCHEIFLINGER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
City: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)856-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SEQ. ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid



NAME: Quine, Jonathan A.  
 REGISTRATION NUMBER: P-41,261  
 REFERENCE/DOCKET NUMBER: 015280-113100US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 9636 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 149..9556  
 US-08-323-170B-1

Query Match 4.6%; Score 47.8; DB 1; Length 9636;  
 Best Local Similarity 49.0%; Pred. No. 0.0063;  
 Matches 127; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

QY 460 ATGATATGACCAAGTCATGATTTACTACAAAGAGCTTAACCAAGATGCTCTAAG 519  
 DB 6941 AGTGATTTAGGAAGATGACATATTTATTTATTTCAATTAAGATTAATAATATTAAT 7000  
 QY 520 TTAAGACATCGACAGCAGCTGTTAACTGGAAGACGTACAGTTAGATGATCCA 579  
 DB 7001 TTAAGACCAAGATGTTTATTAATTAAGATGAAATTTAGAGAAATTAACAAATATA 7060  
 QY 580 CAAATCTACACGATTTACAGATGAAACCCATTAAGATATTTGCTAAGTCT 639  
 DB 7061 CAACATATATACAGATATCAATATTTACAGATGCTTCAATTTCAAAATTTAATTA 7120  
 QY 640 GGGATATCTAATTAAGCAATTTTCAAGATGATCAACCAATTTGATATAT 699  
 DB 7121 GGTATCTACCAATTAATTTAAATAATCTATTTCTACAGCATATGCTAAGTACCAT 7180  
 QY 700 ACTGAATCCCAATTTAA 718  
 DB 7181 ACCTTAATCTATATTA 7199

RESULT 13  
 US-09-286-691-11  
 : Sequence 11, Application US/09286691  
 : Patent No. 6190189  
 : GENERAL INFORMATION:  
 : APPLICANT: Li, Xin-Liang  
 : APPLICANT: Ljungdahl, Lars G.  
 : APPLICANT: Chen, Huizhong  
 : TITLE OF INVENTION: Cellulases and Coding Sequences  
 : FILE REFERENCE: 42-96  
 : CURRENT APPLICATION NUMBER: US/09/286,691  
 : CURRENT FILING DATE: 1999-04-05  
 : EARLIER APPLICATION NUMBER: US 60/027,883  
 : EARLIER FILING DATE: 1996-10-04  
 : EARLIER APPLICATION NUMBER: PCT US97/18008  
 : EARLIER FILING DATE: 1997-10-03  
 : NUMBER OF SEQ ID NOS: 29  
 : SOFTWARE: PatentIn Ver. 2.0  
 : SEQ ID NO 11  
 : LENGTH: 1826  
 : TYPE: DNA  
 : ORGANISM: Ophiomyces sp. PC-2  
 : FEATURE:  
 : NAME/KEY: CDS  
 : LOCATION: (69)..(1481)  
 : US-09-286-691-11

Query Match 4.6%; Score 47.2; DB 4; Length 1826;

Best Local Similarity 46.8%; Pred. No. 0.0053;  
 Matches 148; Conservative 0; Mismatches 168; Indels 0; Gaps 0;

QY 535 GACGAGCTGTACAGGAGAGCTACGTTGAATTAAGATGATCCCAATATCTACAGCT 594  
 DB 1444 gatgggggttgaaatgtaattggttgatattatatttaacttcaatattttaa 1503  
 QY 595 ATTCAGATGAGAAAACCCCTATTAAGTATTTGCTAGCTGGAAATTTGATTTT 654  
 DB 1504 caaacataataattatttagtaataataaagaataaatttttaaaaaatatatt 1563  
 QY 655 AATCAGCAATTTATCAAGATGATCAACACCAATTTGGATTTATACGAAATCCAAT 714  
 DB 1564 tatataatagttataataataataataataatagaattactatagttatataga 1623  
 QY 715 TTAACAGCAATCAACACATATTTGAATTTATTTACTTGAGCTGTGATTTAACACA 774  
 DB 1624 tatatacatacaaaagtaaaaaaatttaaaattttagtattgataaaatttataa 1683  
 QY 775 ATCTTCACAAATTTATTAATAAGAGAGATTGGAAGCTTGCTAGTCGAGGACGTCAC 834  
 DB 1684 aagtttaataatgtaaaaaaataatataaacatttggatgatttgcatacaaga 1743  
 QY 835 ACTACTTCAGATTTCT 850  
 DB 1744 aataataataaataact 1759

RESULT 14  
 US-08-465-795-1  
 : Sequence 1, Application US/08465795  
 : Patent No. 5589355  
 : GENERAL INFORMATION:  
 : APPLICANT: Koizumi, Satoshi  
 : APPLICANT: Yonefani, Yoshiyuki  
 : APPLICANT: Teshiba, Sadao  
 : TITLE OF INVENTION: A Process for Producing Riboflavin  
 : NUMBER OF SEQUENCES: 2  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Finnegan, Henderson, Farbow, Garrett &  
 : STREET: 1300 I Street, N.W., Suite 700  
 : CITY: Washington  
 : STATE: DC  
 : COUNTRY: USA  
 : ZIP: 20005-3315  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: PatentIn Release #1.0, Version #1.25  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/465,795  
 : FILING DATE:  
 : CLASSIFICATION: 435  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 08/161,394  
 : FILING DATE: 06-DEC-1993  
 : APPLICATION NUMBER: JP 326578/1992  
 : FILING DATE: 07-DEC-1992  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Gaybrick, Robert J.  
 : REGISTRATION NUMBER: 27,890  
 : REFERENCE/DOCKET NUMBER: 04853.0008-00000  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 202-408-4000  
 : TELEFAX: 202-408-4400  
 : INFORMATION FOR SEQ ID NO: 1:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 5589 base pairs  
 : TYPE: nucleic acid  
 : STRANDEDNESS: single  
 : TOPOLOGY: linear



```

? MOLECULE TYPE: DNA (genomic)
?
? HYPOTHETICAL: NO
?
? ANTI-SENSE: NO
?
? ORIGINAL SOURCE:
?   ORGANISM: Corynebacterium ammoniagenes
?
US-08-465-795-1

```

Query Match	4.58;	Score 46.4;	DB 1;	Length 5589;
Best Local Similarity	47.18;	Pred. No. 0.012;		
Matches 276;	Conservative 0;	Mismatches 301;	Indels 9;	Gaps 4

OY	31	GTACAAAGCTCAAAAGAGGTAAATCCACCCGCTGGGGCGTTGCTACT---	-AATGAAGT	87
Db	672	GTGGCGGGGACTTACCACGCCGAATCCACCGGTGGGTAAATTATTTTCGACTTCGGGT		731
OY	88	AGGATGTGGTATTTGGTGTGCACATTGGAAAAAGTGACAAAGCATCGGAGGTTCCAGCA		147
Db	732	GAGATTTGGGGCACCGGTGGGACTACAGCCGGTGGGGGGGTGCACGACAGAAAGTCAAGCT		791
OY	148	CTTATATGCGACAAACAAAATGCTGAAGGTGGCAGCATTTATATTTACGTTAGACCATGT		207
Db	792	CTACCCGATGCGCGGGGCAAGACCAGGAAGGCGCTACCGCGGTGGTGCAGCTGGAGCGGTGC		851
OY	208	AGTATTTTTGGTTTCAACACCAACCCGTGTGTAAACAAAATTATGATTTGTAAAGTACGAAA		267
Db	852	CGGATACCGCGCGCACGGAGCGTGACGACGAGCTTTATTTGAACGCGGCAATCAAGAT		911
OY	268	GTAGTATACGCAACMAAAAGAC---AATTGCTTTAGACACACATGATGATAGACGTTACGG		324
Db	912	GTTCTTTTCTTACACCTCGAAGCCGAATCCCAATGGCGGGGGGGAACAGAGGTCTTGC		971
OY	325	GCTACGCGTATTGAGGT--TGAATGCGTTTATGATGACGGGCATACAAATTATACCAAG		382
Db	972	GATCCCGGCAATCAACGTCGTCAGCTGCCAGCCCGGAGGGGTATACGGATGCGCTATT		1031
OY	383	ACTTTTTAAAGCAAAAGCA--AGCACTGCCCAATTAAGTGAAGTATTCGCAAGT		441
Db	1032	CCGTTGGTTGAAGTCCGTCGACGTGCGACGTCCCGACGTTTACGTTGAATTTGGCCAACT		1091
OY	442	TTGATGCGTAAACMAAGCAATGATTAATGACAAAGTCATATGATTTACTTAACAAAGTGTT		501
Db	1092	ATCGATGCGCTTTACTTGGCAGCGCGGATGCCAGCACAGCCAGTGGATTCATTCATCGGCACATGGCAAG		1151
OY	502	AAACAAAGTGTCTTAATGAATTAAGCATCGACACGACGCAAGTGTAACTGGAAGACGTACA		561
Db	1152	CGTGACTATGTCACGCTGCACGTGAACACCGGATGACCATTTATCATTCGGCACATGGCAAG		1211
OY	562	GTTGAATTTAGATGATCCACAATATACAGCATGATTTCAAGATGCA		607
Db	1212	GCCTTGATTTGATTAATTCATGCTTATGCCCCGCTATTCACAGCGCA		1257

RESULT 15  
 US-08-487-8268-13  
 : Sequence 13, Application US/08487826B  
 : Patent No. 5993827  
 :  
 : GENERAL INFORMATION:  
 :  
 : APPLICANT: Sim, Kim L.  
 :  
 : APPLICANT: Chitnis, Chetan  
 :  
 : APPLICANT: Miller, Louis H.  
 :  
 : APPLICANT: Peterson, David S.  
 :  
 : APPLICANT: Su, Xin-zhaun  
 :  
 : APPLICANT: Wellem, Thomas E.  
 :  
 : TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
 : TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE  
 : BINDING PROTEINS  
 :  
 : NUMBER OF SEQUENCES: 45  
 :  
 : CORRESPONDENCE ADDRESS:  
 :

```

1 ZIP: 92660
2
3 COMPUTER READABLE FORM:
4
5 MEDIUM TYPE: Floppy disk
6
7 COMPUTER: IBM PC compatible
8
9 OPERATING SYSTEM: PC-DOS/MS-DOS
10
11 SOFTWARE: PatentIn Release #1.0, Version #1.25
12
13 CURRENT APPLICATION DATA:
14
15 APPLICATION NUMBER: US/08/487,826B
16
17 FILING DATE: 10-SEP-1993
18
19 CLASSIFICATION: 435
20
21 ATTORNEY/AGENT INFORMATION:
22
23 NAME: Israel, N. Ned
24
25 REGISTRATION NUMBER: 29,655
26
27 REFERENCE/DOCKET NUMBER: NH121.001CP1
28
29 TELECOMMUNICATION INFORMATION:
30
31 TELEPHONE: (619) 235-8550
32
33 TELEFAX: (619) 235-0176
34
35 INFORMATION FOR SEQ ID NO: 13:
36
37 SEQUENCE CHARACTERISTICS:
38
39 LENGTH: 19124 base pairs
40
41 TYPE: nucleic acid
42
43 STRANDEDNESS: single
44
45 TOPOLOGY: linear
46
47 MOLECULE TYPE: cdna
48
49 HYPOTHETICAL: NO
50
51 ANTI-SENSE: NO
52
53 US-08-487-826B-13

```

	Query Match	4.2%	Score 43.4	DB 2	Length 19124	
	Best Local Similarity	46.0%	Pred. No. 0.096			
	Matches 188	Conservative 0	Mismatches 216	Indels 5	Gaps 1	
OY	330 TAAAGCAAAAGCAAAGCACTGCCAACAATTCACGTGAAAAGTACTGTGCAGTTTGATNGC	449				
Db	15493 TAAAAAAGAAAAAAAAAAAAAAAAAGGAGAAAAATTTTTTAAAAAATATAAAAAATTATTA	15552				
OY	450 TAAACACCGAATGSATTAATGGACAAGTCATTCGATTTACTAACCAAGAAGCTTAACACGA	509				
Db	15553 AATATTAATTTTGTATAGCATTAATAAATGAAATGAAATTTATCCAAAAAATTTAAAAA	15612				
OY	510 TGCTATATAGTTTAAGACATCGACACAGCGCAGTGTAACTGGAAGACGTACAGTTGATTT	569				
Db	15613 ATTTTATATTAATAAAAAAATGATTTATAAAAA-----AATTAATAAACAAAGAAGAAAAA	15667				
OY	570 AGATATCCACAATTTACTACACGRTTCAGATGSAAGAAAAACCCTATTAAGTATATT	629				
Db	15668 AAAAAACATTAATAAAAAAAAAAATTAATATTCATTAACAAAAACAAAAAATAATATA	15727				
OY	630 GCTCAAGCTCGGGAATTTTCATTTTAATCAGCAAAATTTTCAAGATGAATCAACACCANT	689				
Db	15728 TTAAAAATAAAAATATATATTCATTAATAATTAATAAAAAATTTAAAAAATGTTAAAAA	15787				
OY	690 TTGGATATATACGAAATCCAATTTTAAACAAGCAATCAACACATATTGGAATTTATT	749				
Db	15788 ATATTCATTAATAAATTAATAAAAAATTTTATTAATAAAAAAATAATTAATAAATAATTTA	15847				
OY	750 CTGGAAGCTTGATTTAAACAAGCATTCCTCACATTTATTAATAAGA	798				
Db	15848 ATTAAATTAATAAATAATTAATTAATAAATAATTTTAAATTAATAAAAAA	15896				

Search completed: May 7, 2001, 12:12:52  
Job time: 5648 sec

**THIS PAGE BLANK (USPTO)**